

Case Study 1: An Evidence-Based Practice Review Report

Theme: The evidence base for a single intervention, or particular type of intervention, implemented by parents.

Is parent-led Dialogic Reading an effective intervention for improving vocabulary knowledge of children learning English as an additional language in the UK?

Summary

Dialogic reading (DR), invented by Whitehurst et al. (1988), is an intervention which encourages children's active engagement in shared reading (Urbani, 2020). It is primarily implemented by caregivers within homes and by educators in school settings (Towson et al., 2017).

This systematic literature review aims to evaluate the effectiveness of parent-led DR at improving vocabulary knowledge of children with English as an additional language (EAL) in the UK. A systematic literature search identified five studies that met the inclusion criteria, which were then assessed using the Weight of Evidence (WoE) framework (Gough, 2007). One study was given a rating of 'high' and four studies 'medium'. The effect sizes were calculated and it was found that despite seeing significant outcomes in three of the studies, the effect of parent-led DR on improving vocabulary knowledge of EAL children was small. Generalisability of the findings of these studies may be limited because of where the studies were conducted and

what language was used in the implementation of DR. Recommendations for future research are suggested.

Introduction

Dialogic Reading

Shared interactive reading, a commonly known practice and broadly used term, includes several interventions that aim to increase children's participation in book reading through strategies like child-centeredness, expanding on children's utterances, responding actively, pausing and assessing the children's responses (Hemmeter & Kaiser, 1994). Dialogic Reading (DR), developed by Whitehurst et al. (1988), is a shared interactive reading intervention where an adult uses specific question prompts to motivate the child to talk while they read (Towson et al., 2017). The child becomes the storyteller and parent becomes the listener who assists, facilitates, expands and responds to what the child says during book reading (Pillinger & Vardy, 2022).

DR is a systematic framework developed to support adults in maximising the benefits of shared reading (Pillinger & Vardy, 2022). There are two DR techniques, commonly known by their acronyms – PEER and CROWD. PEER stands for Prompt, Evaluate, Expand, and Repeat. In practice, applying the PEER strategy means the adult prompts the child to name an object on a page or verbally answer their questions about a character in the story, the adult then evaluates the correctness of the child's response, expands on that response by adding some more words and invites the child to repeat the adult's expanded response (Blom-Hoffman et al., 2005). CROWD stands for Completion, Recall, Open-ended, Wh-questions and Distancing. What applying the CROWD strategy looks like in practice

involves the adult asking the child a series of questions: to complete a phrase or sentence, to recall the details of the characters or events in the story, to describe what is happening in a picture, to name the object or action in a picture that the adult is pointing at, and something about the child's own life that is also related to the story (Pillinger & Wood, 2013).

Vocabulary

By talking while reading, it is believed that children's oral language development will be optimised (Towson et al., 2017). Zevenbergen et al. (2003) found DR to be associated with enhanced expressive language and emergent literacy skills. Arnold et al. (1994) and Dale et al. (1996) also found children of parents trained to use DR through videotaped instructions used a higher number of words than the controls in the parent-child interactions. In terms of oral vocabulary, research showed both expressive and receptive vocabulary are associated with competence in reading from Year 4 onwards (Storch & Whitehurst, 2002). Oral vocabulary helps a child to understand print vocabulary when they start becoming literate, and increased print vocabulary knowledge makes their reading more automatic over time (Morgan & Meier, 2008). It was suggested that systematic intervention in preschool years is crucial in improving children's vocabulary, which in turn helps their reading in primary school (Opel et al., 2009). However, despite the wealth of literature around DR, few studies reported the effect size of DR interventions (Towson et al., 2017). It was concluded that the true

educational value of DR can only be confirmed if more studies report effect sizes (Pillinger & Vardy, 2022).

English as an Additional Language

Although there has been a systematic review on the impact of DR on children's literacy (and non-literacy) skills (Pillinger & Vardy, 2022), the impact it has on children whose first language is not English has not yet been explored. According to statistics from the Department for Education (2022), there are over 1.6 million English-as-an-Additional-Language (EAL) students in the UK. EAL learners are defined as those who have been 'exposed to a language at home that is known or believed to be other than English' (Department for Education, 2019).

When compared with monolingual English-speaking same-aged peers, EAL children were found to have an English vocabulary deficit (Mahon & Crutchley, 2006). Moreover, the major factor for EAL children's academic underperformance was found to be difficulties with reading comprehension (Murphy & Unthiah, 2015). As mentioned above, vocabulary knowledge is a key predictor of reading comprehension, and interventions that target vocabulary knowledge is vital in preschool years. Therefore, having noticed a gap in literature around DR improving vocabulary knowledge of EAL children, this review attempts to answer the question around the effectiveness of DR at improving vocabulary knowledge of EAL children.

While parental involvement in a child's reading and learning is often encouraged, it is important to consider that parents may also be new to English or have limited English proficiency. Hence, shared interactive reading practices in English may not be as simple compared to families where English is spoken as their first language. DR in English may be challenging for EAL parents and children due to their limited English proficiency, low self-efficacy in English, and linguistic differences between their first language and the English language (Chow et al., 2009). Practical considerations around EAL parents conducting DR in English lead one to consider the effectiveness of DR as an intervention to improve EAL children's vocabulary knowledge. It is believed that this review will be of great interest and significance to preschools and Educational Psychologists who may wish to recommend parent-led dialogic reading as an intervention to support these children. Therefore, to reiterate, the research question of this review is: *Is parent-led Dialogic Reading an effective intervention for improving vocabulary knowledge of children learning English as an additional language in the UK?*

Critical review of the evidence

Systematic Literature Search

A comprehensive search was conducted on 5th February 2023 using three databases, which produced 20 results. The search terms “dialogic reading” and “interactive shared reading” were used, which were the same search terms used in the most recent systematic review of DR (Pillinger & Vardy, 2022). All search terms are shown in Table 1 while the full literature search process can be visualised using the flow diagram in Figure 1. After duplicates were removed ($n = 7$), the articles were screened at title and abstract level using the inclusion and exclusion criteria (see Table 2). Seven articles were then removed, resulting in six screened at full text level and five selected for this review. Full reference for the excluded article at full text level and the corresponding rationale for exclusion can be found in Appendix A. Key information about the five studies that are relevant to the review question are summarised in Table 4.

Table 1

Search Terms Used

Data base	Search terms
PsycInfo	(“dialogic reading” OR “interactive shared reading”) AND (parent* OR famil* OR home* OR mother* OR father* OR carer* OR caregiver*) AND ("english as an additional language" OR "english as a foreign language" OR "english as a second language" OR "english language learner*" OR "EAL" OR "EFL" OR "ESL" OR "ELL")
ERIC	(“dialogic reading” OR “interactive shared reading”) AND (parent* OR famil* OR home* OR mother* OR father* OR carer* OR caregiver*) AND ("english as an additional language" OR "english as a foreign language" OR "english as a second language" OR "english language learner*" OR "EAL" OR "EFL" OR "ESL" OR "ELL")
Web of Science	(“dialogic reading” OR “interactive shared reading”) AND (parent* OR famil* OR home* OR mother* OR father* OR carer* OR caregiver*) AND ("english as an additional language" OR "english as a foreign language" OR "english as a second language" OR "english language learner*" OR "EAL" OR "EFL" OR "ESL" OR "ELL")

Note. Truncation (*) was used to ensure that multiple endings of root words would be detected.

Figure 1

Flow diagram of the literature search and screening process

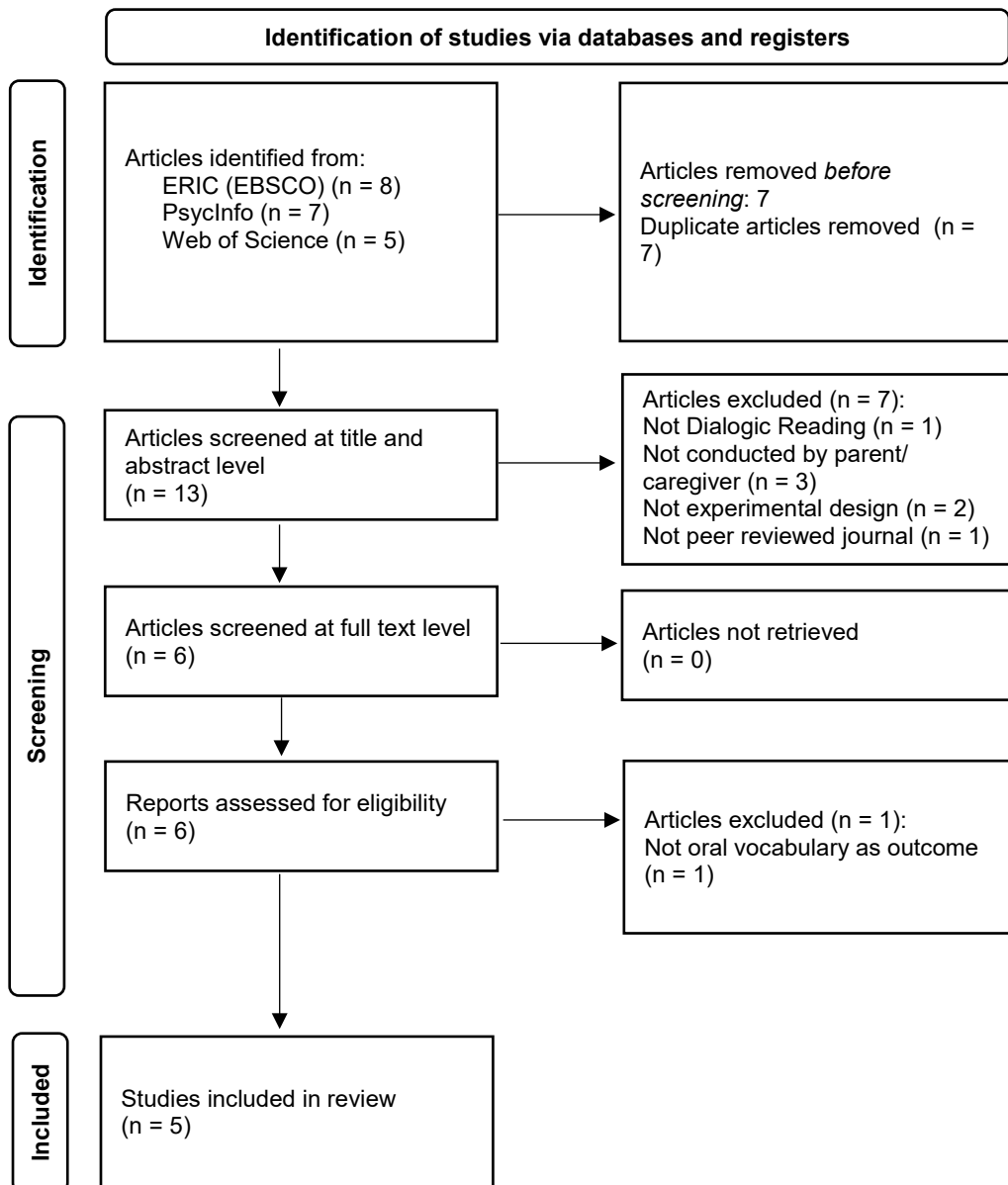


Table 2

Inclusion and exclusion criteria with rationale

	Inclusion Criteria	Exclusion Criteria	Rationale
1 Intervention	Dialogic reading	Not dialogic reading	To investigate whether the results are due to the intervention
2 Study Design	Experimental design	Not experimental design	To compare the effectiveness of the intervention quantitatively
3 Language	Studies written in English	Studies not written in English	There are no translation services available
4 Published date	Studies published since 2000	Studies published before 2000	Studies from the 2000s are considered more recent and have higher generalisability
5 Type of publication	Peer reviewed	Not peer reviewed	Peer reviewed journals ensure high quality research due to screening for validity and quality control before publication
6 Participants	EAL children	Not EAL children	The purpose of this review is to investigate the effectiveness of the intervention on EAL children
7 Interventionists	Parents or caregivers of EAL children	Not parents or caregivers of EAL children	The purpose of this review is to investigate the effectiveness of the intervention implemented by parents or caregivers
8 Outcome measure	Vocabulary	Not vocabulary	The purpose of this review is to investigate the effectiveness of the intervention on vocabulary only

Table 3*List of Included Studies*

Included Studies
Brannon, D., & Dauksas, L. (2012). Increasing the Expressive Vocabulary of Young Children Learning English as a Second Language Through Parent Involvement. <i>Procedia - Social and Behavioral Sciences</i> , 69, 1324–1331. https://doi.org/10.1016/j.sbspro.2012.12.069
Brannon, D., & Dauksas, L. (2014). <i>The Effectiveness of Dialogic Reading in Increasing English Language Learning Preschool Children's Expressive Language</i> . https://doi.org/10.4225/03/5817d8a638fe4
Chow, B. W.-Y., McBride-Chang, C., & Cheung, H. (2009). Parent-child reading in English as a second language: Effects on language and literacy development of Chinese kindergarteners. <i>Journal of Research in Reading</i> , 33(3), 284–301. https://doi.org/10.1111/j.1467-9817.2009.01414.x
Petchprasert, A. (2014). The Influence of Parents' Backgrounds, Beliefs about English Learning, and a Dialogic Reading Program on Thai Kindergarteners' English Lexical Development. <i>English Language Teaching</i> , 7(3). https://doi.org/10.5539/elt.v7n3p50
Yang, D., Xia, C., Collins, P., & Warschauer, M. (2022). The role of bilingual discussion prompts in shared E-book reading. <i>Computers & Education</i> , 190, 104622. https://doi.org/10.1016/j.compedu.2022.104622

Table 4

Mapping the Field

Study	Country	Participants	Design	Implementation	Measures related to this review	Relevant significant outcomes
Brannon and Dauksas (2012)	America	30 preschool children from low-income Hispanic families (between ages 3 and 5)	Quasi-experimental Two groups, pre-post (DR and control condition)	DR strategy: CAR, 1, 2, 3 Tell Me What You See, modelling Duration: 10 weeks Frequency: unknown Materials: a random set of 5 picture books in English and Spanish DR: reading the books while giving prompts Control: reading the same books without giving prompts	The picture naming section of the Individual Growth Developmental Indicators (IGDI) test: Children were shown pictures on individual cards and asked to name as many picture objects as they could within 1 minute. (test of expressive vocabulary)	The DR group had significantly more words correct (M=13.29, $p < .01$) in post-test compared to the control group (M=12.38).

Study	Country	Participants	Design	Implementation	Measures related to this review	Relevant significant outcomes
Brannon and Dauksas (2014)	America	41 preschool children from low-income Hispanic families (between ages 3 and 5)	Quasi-experimental Two groups, pre-post (DR and control condition)	DR strategy: CAR, 1, 2, 3 Tell Me What You See, modelling Duration: 10 weeks Frequency: unknown Materials: a random set of 5 picture books in English and Spanish DR: reading the books while giving prompts Control: reading the same books without giving prompts	The picture-naming section of the Individual Growth Developmental Indicators (IGDI) test Children were shown pictures on individual cards and asked to name as many picture objects as they could within 1 minute. (test of expressive vocabulary)	The DR group had significantly more words correct (M=14.32, p<.01) in post-test compared to the control group (M=12.48).

Study	Country	Participants	Design	Implementation	Measures related to this review	Relevant significant outcomes
Chow et al. (2009)	China	51 EAL third-year kindergartners (aged between 51 and 71 months)	Randomised controlled trials Three groups, pre-post (DR, typical reading (TR) and control condition)	DR strategy: PEER Duration: 12 weeks Frequency: twice a week, 20 minutes each Materials: 12 English books DR: parents used the PEER strategy when reading the book TR: used the same books without the DR strategy Control: no reading materials were provided until post-test, read usual reading materials	Peabody Picture Vocabulary Test – Third Edition (PPVT-III; Dunn & Dunn, 1997) with 204 items: Experimenter provided the words orally and for each word, children had to select the correct illustration out of four illustrations. (test of receptive vocabulary)	No significant differences in scores were found between the groups.

Study	Country	Participants	Design	Implementation	Measures related to this review	Relevant significant outcomes
Petchprasert (2014)	Thailand	54 EAL kindergartners (between ages 3 and 6)	Quasi-experimental one group, pre-post	DR strategy: PEER Duration: 4 weeks Frequency: 7 days per week, 30-40 minutes per day Materials: 4 Aesop fables Parents read one Aesop fables provided to them each week using the PEER strategy	14-picture-vocabulary test developed by the researcher according to the picture vocabulary in an Aesop book provided to parents: Children had to select the correct picture from four pictures for each vocabulary heard. (test of receptive vocabulary)	The post-test scores were significantly higher (M=8.26, p=.00) than pre-test scores (M=3.70).

Study	Country	Participants	Design	Implementation	Measures related to this review	Relevant significant outcomes
Yang et al. (2022)	China	107 EAL children (between ages 3 and 7)	Randomised controlled trials (DR condition and control)	DR strategy: CROWD Duration: 2 shared reading sessions Materials: 1 e-book "The Story of an Orange Oakleaf." DR: using discussion prompts embedded in the e-book, prompts were designed following the CROWD strategy Control: using the same e-book without discussion prompts	50-item English Story Vocabulary Test: Receptive Vocabulary subtest (25 items) – children listened to an English word or phrase and had to select the correct picture out of four pictures. Expressive Vocabulary subtest (25 items) – children were shown 4 pictures with 5-8 story-related items on each picture, they then had to name the item in English when the test administrator pointed to it with their mouse.	No significant differences in scores were found between the groups.

Weight of Evidence (WoE)

This review uses the framework by Gough (2007) to appraise the five studies. Gough's framework includes three dimensions – Weight of Evidence A (WoE A), Weight of Evidence B (WoE B) and Weight of Evidence C (WoE C), and an average of the three aforementioned criteria is taken to give an overall numbered rating, i.e., Weight of Evidence D (WoE D). WoE A evaluates the quality of each study according to an agreed criteria. This review uses the quality indicators of Gersten et al. (2005) to review all studies as two studies are randomised controlled trials (RCTs) and three quasi-experimental. WoE B evaluates the relevance of the methodological design of each study to the review question and WoE C considers how relevant the focus of each study is to the review question. The details of the WoE A coding protocol (Gersten et al., 2005) with individual ratings of each study can be found in Appendix B.

WoE A – Methodological quality of the research

Only one coding protocol, i.e., Gersten et al. (2005), is used to review RCTs and quasi-experimental studies. The total score is calculated by adding together the scores from the essential criteria (9 or more = 1, less than 9 = 0) and desirable criteria (4 or more = 2, 1 to 3 = 1, 0 = 0). Table 5 shows the rating criteria of WoE A based on guidance by Gersten et al. (2005). High quality studies (WoE A score of 3) meet 9 essential quality indicators and at least 4 desirable quality indicators; acceptable quality studies (WoE A score of 2) meet 9 essential quality indicators and at least 1 desirable quality indicator. Studies with a WoE A score of 1 are considered poor quality. Table 6 displays the judgement ratings of all five studies.

Table 5

WoE A Criteria for Group Experimental and Quasi-experimental Studies

Criteria	WoE A Rating
9 essential criteria met and at least 4 desirable criteria met	3 (High)
9 essential criteria met and at least 1 desirable criteria met	2 (Medium)
Less than 9 essential criteria met	1 (Low)

Note. WoE A criteria is based on the guidance by Gersten et al. (2005)

Table 6

WoE A Judgement Ratings

Study	Essential Criteria	Desirable Criteria	WoE A
Brannon and Dauksas (2012)	7	2	1 (Low)
Brannon and Dauksas (2014)	7	3	1 (Low)
Chow et al. (2009)	9	2	2 (Medium)
Petchprasert (2014)	4	3	1 (Low)
Yang et al. (2022)	9	2	2 (Medium)

Note. WoE A score of 1 = Low; 2 = Medium; 3 = High.

WoE B – Methodological relevance

WoE B is assessed using the hierarchy from Petticrew and Roberts (2003).

Table 7 shows the rating criteria of WoE B, with 3 being the highest, and 1

lowest. Table 8 displays WoE B ratings of each study.

Table 7

WoE B Criteria

Criteria	WoE B Rating
Randomised controlled trials	3 (High)
Cohort studies, quasi-experimental studies, single case experimental designs	2 (Medium)
Qualitative research, survey, case control, non-experimental evaluation	1 (Low)

Note. WoE B criteria is based on the hierarchy from Petticrew and Roberts (2003)

Table 8

WoE B Judgement Ratings

Study	Weight of Evidence B
Brannon and Dauksas (2012)	2 (Medium)
Brannon and Dauksas (2014)	2 (Medium)
Chow et al. (2009)	3 (High)
Petchprasert (2014)	2 (Medium)
Yang et al. (2022)	3 (High)

Note. WoE B score of 1 = Low; 2 = Medium; 3 = High.

WoE C – Topic relevance

WoE C criteria were developed to evaluate the relevance of the studies to the review question. The criteria with rationale can be found in Table 9. The final WoE C score of each study is determined by calculating the average of the judgment ratings from criterion A to D. Table 10 displays the WoE C ratings of each study.

Table 9

WoE C Criteria and Rationale

Criteria	Ratings	Rationale
A Intervention method	3. DR principles designed by Whitehurst and colleagues (Whitehurst et al., 1988) and appropriate materials are used	This is to ensure DR is used in its 'pure' form as designed by Whitehurst and colleagues (Whitehurst et al., 1988)
	2. DR does not follow the principles designed by Whitehurst and colleagues (Whitehurst et al., 1988)	
	1. Unclear if DR is used/ unclear how the intervention is carried out	
B Intervention language	3. Only English is used in the intervention	The results will be more generalisable to UK settings if the intervention uses English intended for building a child's English vocabulary rather than the first language
	2. Both English and the child's first language are used in the intervention	
	1. Unclear if English is used/ only the child's first language is used in the intervention	
C Participants' EAL status	3. Participants are clearly defined as EAL children and a quantifiable measure of their English language exposure/ usage at home is provided	This is to ensure the intervention has an effect on the intended group of participants
	2. Participants are clearly defined as EAL children but no quantifiable measure of their English language exposure/ usage at home is provided	
	1. Unclear if participants are EAL children	

Criteria	Ratings	Rationale
D Country of study	<ol style="list-style-type: none">3. Study is carried out in the UK2. Study is carried out in a country other than the UK where English is the official first language1. Study is carried out in a country where English is not the official first language	The results will be more generalisable to UK settings where the EAL families' first language is only spoken by the minority population

Table 10

WoE C Judgement Ratings

Study	Intervention method	Intervention language	Participants' EAL status	Country of Study	WoE C
Brannon and Dauksas (2012)	2	1	2	2	1.75
Brannon and Dauksas (2014)	2	1	2	2	1.75
Chow et al. (2009)	3	3	3	1	2.5
Petchprasert (2014)	3	3	2	1	2.25
Yang et al. (2022)	3	2	2	1	2

Table 11

Summary of WoE A, B, C and D

Study	WoE A	WoE B	WoE C	WoE D
Brannon and Dauksas (2012)	1	2	1.75	1.58 (Medium)
Brannon and Dauksas (2014)	1	2	1.75	1.58 (Medium)
Chow et al. (2009)	2	3	2.5	2.5 (High)
Petchprasert (2014)	1	2	2.25	1.75 (Medium)
Yang et al. (2022)	2	3	2	2.33 (Medium)

Note. WoE D of each study is calculated by taking the average of WoE A, B and C. WoE D score of 1.4 or less = Low; 1.5-2.4 = Medium; 2.5-3 = High.

Participants

In total, 283 participants from three countries across five studies were included in this review. The ages of EAL children ranged from 3 to 7 years old. In all but one study (i.e., Yang et al., 2022), children were in preschool and within the age range of 3 to 6. The children's education levels were not reported in Yang et al. (2022). In the studies, it was clearly stated that participants in Brannon and Dauksas (2012; 2014) spoke Spanish as their primary language in the home; in Petchprasert (2014), all participants spoke Thai as their primary language; and in Yang et al. (2022), all children had Chinese as their first language, same for their parents. In Chow et al. (2009), from the descriptive statistics provided in the demographic measures, it was clear that more time was spent on speaking in Chinese than English in the parent-child daily conversation time rated by the parent across all three conditions. Therefore, their primary language was Chinese. All studies were given high scores for the quality indicator in the Gersten et al. (2005) coding protocol for WoE A regarding whether sufficient information was provided to ascertain the participants demonstrated the difficulties presented, defined as EAL in this review. Chow et al. (2009) specifically scored the highest (score of 3) in the participants' EAL status criterion in WoE C as a quantifiable measure of their first language and English language usage was provided.

In Brannon and Dauksas (2012; 2014), the initial survey was used and found there were no significant differences between the experimental and control groups. Factors that could potentially give an advantage to one group over the other were considered – in Brannon and Dauksas (2012), these included

children's participation in after-school programmes, how many adults and children were living in their home, how often children saw a parent read at home, number of books in the home and visits to the library. In Brannon and Dauksas (2014), there were no significant differences between the experimental and control groups in respect of parental education, home language, how often children saw a parent read at home, number of books in the home and visits to the library. Hence, both studies were given high scores for the quality indicator about whether appropriate procedures were used to make the characteristics of participants more likely comparable across conditions in WoE A.

Study Design

Four studies in this review had a control group in their design with clear descriptions (Brannon & Dauksas, 2012; 2014; Chow et al., 2009; Yang et al., 2022). In Brannon and Dauksas (2012; 2014), parents in the control group read aloud the same books as the DR group to their children without any further instruction. In Chow et al. (2009), children in the control group were given the same set of storybooks and DR training materials only post-test. There was also another comparison group, which was the Typical Reading (TR) group, where children were given the same set of storybooks as the DR group but without hints for prompt questions in them. In Yang et al. (2022), parents in the control group read the e-book without discussion prompts with their children. It is important to note that the procedure in the TR group in Chow et al. (2009), compared to its control group, was more

comparable to the control groups in the other three studies. This is because the control group participants in the other three studies were given the same reading materials the same time as the DR group, which was the same for the TR group in Chow et al. (2009), while the control group in this study was given no reading materials until post-test.

Chow et al. (2009) and Yang et al. (2022) were RCTs, hence were rated higher in WoE B (see Table 8) according to the hierarchy from Petticrew and Roberts (2003). Only one coding protocol (Gersten et al., 2005) was used in this review. Gersten et al. (2005) presents quality indicators for both experimental and quasi-experimental studies, which were the research designs of all five studies in this review. Petchprasert (2014) did not have a control group, the low number of essential quality indicators met for this study can be explained by some of them being unknown or unable to code. Moreover, none of the studies measured the maintenance of effects of DR, hence all of them were given low scores for the quality indicator about whether the outcomes were measured beyond an immediate post-test in WoE A.

Intervention implementation

There were variations in how DR was implemented, and how long it ran, across the five studies. Brannon and Dauksas (2012; 2014) used the “Comment and wait, Ask questions and wait, Respond and add more” (CAR) (Washington Research Institute, 1997), and “1, 2, 3 Tell Me What You See” strategies, while Chow et al. (2009) and Petchprasert (2014) used the PEER strategy, and Yang et al. (2022) the CROWD strategy. Both PEER and CROWD were developed by Whitehurst and colleagues (Whitehurst et al., 1988). Therefore, Chow et al., (2009), Petchprasert (2014) and Yang et al. (2022) were rated higher in the intervention method criterion of WoE C.

In terms of programme duration, the implementation of DR varied from two shared reading sessions to twelve weeks, which is a large range. Moreover, the number of reading materials varied greatly. Yang et al. (2022) used only one e-book, while Chow et al. (2009) used twelve physical books. Yang et al. (2022) was the only study in this review that did not use a physical book. The e-book was also the only reading material among all five studies to be bilingual. The use of both the children’s first language and English in text (with only English audio narration) resulted in a medium score in the implementation language criterion of WoE C for this study. As for Brannon and Dauksas, (2012; 2014), since it was unclear whether the intervention language was in the participants’ first language, English or both, these two studies were given low scores in the implementation language criterion of WoE C.

Compared to prompts being provided by parents themselves after being trained to use DR in the other four studies, Yang et al. (2022) had discussion prompts that were provided consistently in the e-book, hence the quality and fidelity of implementation is arguably higher than other studies. However, it is not the same as researchers having actually assessed the fidelity and examined the quality of implementation. In all studies except Yang et al. (2022), parents received training in using DR; among which Chow et al. (2009) involved researchers contacting parents in both experimental groups over the phone to check if they had encountered any problems post-training. Petchprasert (2014) also involved the teachers facilitating and mentoring the parents and children to check if they had encountered any problems. Although steps were taken to monitor the programme, there were no further details or evidence provided that the fidelity or quality of implementation was evaluated. In Brannon and Dauksas (2012; 2014), researchers used the Adult-Child Interactive Reading Inventory (ACIRI) (DeBruin-Parecki, 1999) to analyse their interactions in the videos; researchers in Yang et al. (2022) also selected 20 videos to qualitatively explore the parent-child interactions. Although the implementation process in these three studies were filmed, rather than to evaluate or ensure fidelity, it was one of their measures. Hence, all five studies had low scores for two quality indicators related to fidelity in WoE A.

Measures

The five studies used similar measures to assess children's oral receptive vocabulary, as well as oral expressive vocabulary knowledge. Brannon and Dauksas (2012; 2014) measured expressive vocabulary, while Chow et al. (2009) and Petchprasert (2014) measured receptive vocabulary. Yang et al. (2022) was the only study that measured both expressive and receptive vocabulary. Brannon and Dauksas (2012; 2014) used the picture-naming section of the Individual Growth Developmental Indicators (IGDI), and the 2014 study provided its one-month alternate form reliability coefficients range from $r = .44$ to $.78$ (McConnell et al., 2002). Chow et al. (2009) used Peabody Picture Vocabulary Test – Third Edition (PPVT-III; Dunn & Dunn, 1997). Petchprasert (2014) developed their own picture-vocabulary test according to the picture vocabulary in an Aesop book provided to parents, and provided an overall reliability score ($\alpha = .78$). Yang et al. (2022) also developed their own English Story Vocabulary Test with a reliability score ($\alpha = .97$). Both studies provided internal consistency data, but did not provide test-retest reliability data, while those using standardised tests, i.e., IGDI and PPVT-III did not provide any. Hence, all studies had low scores for that quality indicator in WoE A.

Regarding whether evidence of criterion-related validity and construct validity of the measures were provided, only Brannon and Dauksas (2014) provided evidence, i.e., the results correlated with PPVT-III (Dunn & Dunn, 1997) and Preschool Language Scale (McConnell et al., 2000). Hence, besides

Brannon and Dauksas (2014), other studies had low scores for that quality indicator in WoE A.

Outcomes

Three out of five studies reported statistically significant effects (Brannon & Dauksas, 2012; 2014; Petchprasert 2014). In Brannon and Dauksas (2012), the DR group had significantly more correct words ($M=13.29$, $p<.01$) in post-test than the control group ($M=12.38$). In Brannon and Dauksas (2014), the DR group had significantly more correct words ($M=14.32$, $p<.01$) in post-test than the control group ($M=12.48$). In Petchprasert (2014), post-test scores were significantly higher ($M=8.26$, $p=.00$) than pre-test scores ($M=3.70$). Both Chow et al. (2009) and Yang et al. (2022) reported effect sizes, hence they had high scores for that quality indicator in WoE A. For Brannon and Dauksas, (2012; 2014), Cohen's d values were 0.40 and 0.31 respectively. According to Cohen's (1988) criteria, 0.2 is a small effect size, 0.5 is moderate and 0.8 is large. As can be seen, the effect of DR on improving vocabulary knowledge of children in the studies was quite small.

Petchprasert (2014) had a large effect size ($d=1.00$) using the pooled standard deviation controlling for the intercorrelation of both pre- and post-test groups (Lakens, 2013). However, it is important to note that there was no control group in this study, and the improvement in vocabulary knowledge could happen anyway with maturation, development and continued attendance at the preschool etc. Hence, without the control group as comparison, even with a large effect size, the effect of DR on improving the

children's vocabulary knowledge was unconvincing. Table 12 shows the measures and effect sizes, together with WoE D ratings. It is important to note that the effect size of Chow et al. (2009) in Table 12 was calculated using scores of DR and TR groups, instead of DR and the control group. This is because the control conditions of the other three studies shared the same procedures as TR (where the same reading materials as DR group were used during the programme), rather than control group (where participants received the reading materials post-test). Hence, it was decided that the scores of the TR group would be computed in the calculation of Cohen's *d* for Chow et al. (2009) as a control. In fact, in this study, it was found that by using the same reading materials without applying DR principles, the TR group had a greater gain in English receptive vocabulary than the DR group, hence the effect size was found to be a negative value ($d=-0.14$), suggesting DR not only did not improve vocabulary knowledge of these EAL children, but indeed worsened it. However, when compared to the control group where no reading materials were given until after the programme ended, the families used their usual reading materials, and the effect size was also very small ($d=0.02$), confirming DR had little effect on improving vocabulary knowledge of EAL children.

Table 12

The measures and effect sizes of all studies with WoE D ratings

Study	Sample size	Outcome measure	Analysis	Significance	Effect size (Cohen's d)	WoE D Rating
Brannon and Dauksas (2012)	30	Picture naming section of the Individual Growth Developmental Indicators (IGDI) test	Between-group post-intervention comparison, controlling for pre-post differences	p<.01	0.396 (Small)	1.67 (Medium)
Brannon and Dauksas (2014)	41	Picture naming section of the Individual Growth Developmental Indicators (IGDI) test	Between-group post-intervention comparison, controlling for pre-post differences	p<.01	0.306 (Small)	1.67 (Medium)
Chow et al. (2009)	51	Peabody Picture Vocabulary Test – Third Edition (PPTV-III)	Between-group post-intervention comparison, controlling for pre-post differences	n.s.	-0.138	2.5 (High)
Petchprasert (2014)	54	A 14-picture vocabulary test developed by the researcher	Repeated measures pre-post intervention	p=.00	0.993 (Large)	1.75 (Medium)
Yang et al. (2022)	107	A 50-item English Story Vocabulary Test developed by the researcher	Between-group post-intervention comparison, controlling for pre-post differences	n.s.	0.1212	2.33 (Medium)

Conclusion and Recommendations

Conclusion

This review investigated the effectiveness of parent-led DR at improving vocabulary knowledge of EAL children. Five studies met the criteria and only three found improvements in children's vocabulary knowledge after the implementation of DR, although the effect of DR was found to be small (Brannon & Dauksas, 2012; 2014). One study even found it to negatively impact children's improvement in vocabulary knowledge (Chow et al., 2009), although compared to its control group, the effect size was very minimal, rather than negative. Although Petchprasert (2014) found a large effect size, due to the lack of a control group, it was difficult to ascertain if the children's improvement in vocabulary knowledge was due to DR. Moreover, the findings in Brannon and Dauksas (2012; 2014) had to be viewed with caution, due to the ambiguity of intervention language used, and its effect on English vocabulary knowledge, rather than their first language. Hence, taking all these factors into consideration, this review could not ascertain the effectiveness of parent-led DR on improving vocabulary knowledge of EAL children. It is important to note that although all studies achieved a WoE D rating of 'medium' and above, three studies had low WoE A ratings, which suggests the studies were not well-conducted, and this would impact findings, despite their relevance to the research question. Nonetheless, DR has benefits such as cost-effectiveness of videotaped instructions for DR training, hence may be favoured by some families and have considerable educational significance.

Recommendations

As none of the studies measured DR's effect beyond the immediate post-test, it will be helpful for future research to measure if DR has a lasting impact on children's improvement in oral vocabulary knowledge. Moreover, a control group should be used to ascertain if it was DR that had an impact on any improvements in children's vocabulary knowledge.

To answer the review question better, future research should be conducted in the UK so that findings will be more generalisable to UK EAL families. In the UK, their first language usage is restricted to their home environments only, this is in contrast to some studies conducted in countries where their first language is spoken in almost all settings. Therefore, the importance of English also varies between the UK and these countries.

Future studies should also make clear the intervention language so it will be easier to determine if the study is truly relevant to English vocabulary knowledge acquired during the intervention. Parents' self-report of their confidence and proficiency in English will remain useful. If parents' English proficiency is limited or if they do not feel comfortable speaking in English, there will be difficulties when DR is implemented in the intended English language, which may ultimately affect its effectiveness on improving English vocabulary knowledge of EAL children.

References

- Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology, 86*(2), 235–243. <https://doi.org/10.1037/0022-0663.86.2.235>
- Blom-Hoffman, J., O’Neil-Pirozzi, T. M., & Cutting, J. (2005). Read together, talk together: The acceptability of teaching parents to use dialogic reading strategies via videotaped instruction. *Psychology in the Schools, 43*(1), 71–78. <https://doi.org/10.1002/pits.20130>
- Brannon, D., & Dauksas, L. (2012). Increasing the Expressive Vocabulary of Young Children Learning English as a Second Language Through Parent Involvement. *Procedia - Social and Behavioral Sciences, 69*, 1324–1331. <https://doi.org/10.1016/j.sbspro.2012.12.069>
- Brannon, D., & Dauksas, L. (2014). *The Effectiveness of Dialogic Reading in Increasing English Language Learning Preschool Children’s Expressive Language*. <https://doi.org/10.4225/03/5817d8a638fe4>
- Chow, B. W.-Y., McBride-Chang, C., & Cheung, H. (2009). Parent-child reading in English as a second language: Effects on language and literacy development of Chinese kindergarteners. *Journal of Research in Reading, 33*(3), 284–301. <https://doi.org/10.1111/j.1467-9817.2009.01414.x>
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Journal of the American Statistical Association, 73*(363), 680. <https://doi.org/10.2307/2286629>
- 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. *Topics in Early Childhood Special Education, 16*(2), 213–235. <https://doi.org/10.1177/027112149601600206>
- Department for Education (2019). *Attainment of pupils with English as an additional language* [White paper]. Crown. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/908929/Attainment_of_EAL_pupils.pdf
- Department for Education. (2022). *Schools, Pupils and Their characteristics, Academic Year 2021/22*. Explore-Education-Statistics.service.gov.uk. <https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics>

- DeBruin-Parecki, A. (1999). *Adult/Child Interactive Reading Inventory*. Center for the Improvement of Early Reading Achievement.
<https://files.eric.ed.gov/fulltext/ED447412.pdf>
- Dunn, L.M. & Dunn, L.M. (1997). Peabody picture vocabulary test. (3rd edn). Circle Pines, MN: American Guidance Service.
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional Children*, 71(2), 149–164.
<https://doi.org/10.1177/001440290507100202>
- Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213–228.
<https://doi.org/10.1080/02671520701296189>
- Hemmeter, M. L., & Kaiser, A. P. (1994). Enhanced Milieu Teaching. *Journal of Early Intervention*, 18(3), 269–289.
<https://doi.org/10.1177/105381519401800303>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4(863).
<https://doi.org/10.3389/fpsyg.2013.00863>
- Mahon, M., & Crutchley, A. (2006). Performance of typically-developing school-age children with English as an additional language on the British Picture Vocabulary Scales II. *Child Language Teaching and Therapy*, 22(3), 333–351. <https://doi.org/10.1191/0265659006ct311xx>
- McConnell, S.R., Priest, J.S., Davis, S.D., & McEvoy, M.A. (2002). Best practices in measuring growth and development for preschool children. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., Vol. 2, pp. 1231-1246). Washington, DC : National Association of School Psychologists.
- Morgan, P. L., & Meier, C. R. (2008). Dialogic Reading's Potential to Improve Children's Emergent Literacy Skills and Behavior. *Preventing School Failure : Alternative Education for Children and Youth*, 52(4), 11–16.
<https://doi.org/10.3200/PSFL.52.4.11-16>
- Murphy, V., & Unthiah, A. (2015). *A systematic review of intervention research examining English language and literacy development in children with English as an Additional Language (EAL)*.
<http://www.naldic.org.uk/Resources/NALDIC/Research%20and%20Information/Documents/eal-systematic-review-prof-v-murphy.pdf>

- Opel, A., Ameer, S. S., & Aboud, F. E. (2009). The effect of preschool dialogic reading on vocabulary among rural Bangladeshi children. *International Journal of Educational Research*, 48(1), 12–20. <https://doi.org/10.1016/j.ijer.2009.02.008>
- Petchprasert, A. (2014). The Influence of Parents' Backgrounds, Beliefs about English Learning, and a Dialogic Reading Program on Thai Kindergarteners' English Lexical Development. *English Language Teaching*, 7(3). <https://doi.org/10.5539/elt.v7n3p50>
- Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology & Community Health*, 57(7), 527–529. <https://doi.org/10.1136/jech.57.7.527>
- Pillinger, C., & Vardy, E. J. (2022). The story so far: A systematic review of the dialogic reading literature. *Journal of Research in Reading*. <https://doi.org/10.1111/1467-9817.12407>
- 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. <https://doi.org/10.1177/1468798413491975>
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, 38(6), 934–947. <https://doi.org/10.1037/0012-1649.38.6.934>
- Towson, J. A., Fettig, A., Fleury, V. P., & Abarca, D. L. (2017). Dialogic Reading in Early Childhood Settings: A Summary of the Evidence Base. *Topics in Early Childhood Special Education*, 37(3), 132–146. <https://doi.org/10.1177/0271121417724875>
- Urbani, J. M. (2020). Dialogic Reading: Implementing an Evidence-Based Practice in Complex Classrooms. *TEACHING Exceptional Children*, 52(6), 392–402. <https://doi.org/10.1177/0040059920917694>
- Washington Research Institute (1997). Language is the key: A program for building language and literacy in early childhood [Motion picture]. (Available from Washington Learning Systems, 2212 Queen Anne Avenue North, #726, Seattle, WA 98109)
- Wells, C. S., & Wallack, J. A. (2003). An Instructor's Guide to Understanding Test Reliability. Retrieved August 25, 2013, from <http://testing.wisc.edu/Reliability.pdf>
- Whitehurst, G. J., Falco, F. L., Lonigan, C. J., & Fischel, J. E. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, 24(4), 552–559. <https://doi.org/10.1037/0012-1649.24.4.552>

Yang, D., Xia, C., Collins, P., & Warschauer, M. (2022). The role of bilingual discussion prompts in shared E-book reading. *Computers & Education, 190*, 104622.
<https://doi.org/10.1016/j.compedu.2022.104622>

Zevenbergen, A. A., Whitehurst, G. J., & Zevenbergen, J. A. (2003). Effects of a shared-reading intervention on the inclusion of evaluative devices in narratives of children from low-income families. *Journal of Applied Developmental Psychology, 24*(1), 1–15.
[https://doi.org/10.1016/s0193-3973\(03\)00021-2](https://doi.org/10.1016/s0193-3973(03)00021-2)

Appendices

Appendix A

Table 1

List of Excluded Studies at Full Text Screening level with Rationale

Excluded Studies	Rationale
Huennekens, M. E., & Xu, Y. (2015). Using dialogic reading to enhance emergent literacy skills of young dual language learners. <i>Early Child Development and Care</i> , 186(2), 324–340. https://doi.org/10.1080/03004430.2015.1031125	Not oral vocabulary as outcome

Appendix B1 – WoE A: Methodological quality Coding Protocol

Critical appraisal checklist reference: Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional Children*, 71(2), 149–164.

<https://doi.org/10.1177/001440290507100202>

Study Reference: Brannon, D., & Dauksas, L. (2012). Increasing the Expressive Vocabulary of Young Children Learning English as a Second Language Through Parent Involvement. *Procedia - Social and Behavioral Sciences*, 69, 1324–1331. <https://doi.org/10.1016/j.sbspro.2012.12.069>

Table 2: Scoring for WoE A, based on Gerstens et al (2005) guidelines

	High Quality (Needs to meet 9 essential quality Indicators and at least 4 desirable) Rating = 3	Acceptable Quality (Needs to meet 9 essential quality Indicators and at least 1 desirable) Rating = 2	Low quality (Meets less than 9 of the essential Criteria) Rating = 1	Overall rating (1-3)
Number of essential quality indicators met 7/10			x	1
Number of desirable quality indicators met 2/8				

Essential Quality Indicators**A. Quality Indicators for Describing Participants**

1. Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?
 Yes
 No
 Unknown/Unable to Code

2. Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?
 Yes
 No
 Unknown/Unable to Code

3. Was sufficient information given characterizing the interventions or teachers provided? Did it indicate whether they were comparable across conditions?
 Yes
 No
 Unknown/Unable to Code

B. Quality Indicators for Implementation of the Intervention and Description of Comparison Conditions

4. Was the intervention clearly described and specified?
 Yes
 No
 Unknown/Unable to Code

5. Was the fidelity of implementation described and assessed?
 Yes
 No
 Unknown/Unable to Code

6. Was the nature of services provided in comparison conditions described?
 Yes
 No
 Unknown/Unable to Code

C. Quality Indicators for Outcome Measures

7. Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalized performance?

Yes
 No
 Unknown/Unable to Code

8. Were outcomes for capturing the interventions effect measured at the appropriate times?

Yes
 No
 Unknown/Unable to Code

D. Quality Indicators for Data Analysis

9. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?

Yes
 No
 Unknown/Unable to Code

10. Did the research report include not only inferential statistics but also effect size calculations?

Yes
 No
 Unknown/Unable to Code

Desirable Quality Indicators

11. Was data available on **attrition rates** among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?

Yes
 No
 Unknown/Unable to Code

12. Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally unfamiliar to examinees across study conditions?

- Yes
- No
- Unknown/Unable to Code

13. Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?

- Yes
- No
- Unknown/Unable to Code

14. Was evidence of the criterion-related validity and construct validity of the measures provided?

- Yes
- No
- Unknown/Unable to Code

15. Did the research team assess not only surface features of fidelity implementation but also examine quality of implementation?

- Yes
- No
- Unknown/Unable to Code

16. Was any documentation of the nature of instruction or series provided in comparison conditions?

- Yes
- No
- Unknown/Unable to Code

17. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

- Yes
- No
- Unknown/Unable to Code

18. Were results presented in a clear, coherent fashion?

- Yes
- No
- Unknown/Unable to Code

Overall Rating of Evidence: 3 2 1