Case Study 1: An Evidence-Based Practice Report

Theme: Parent-delivered interventions for children with Special Educational Needs (SEN)

What is the impact of dialogic reading between parent and child on child language development?

1.0 Summary

Dialogic Reading interventions have a longstanding reputation within the field of education as a valid intervention to support children's language development for both neurotypical children and children with learning and developmental disabilities (Nunes et al., 2022). However, despite significant reputation in the educational field, far less research has investigated the impact of Dialogic Reading when delivered by parents. This review aimed to evaluate the impact that Dialogic Reading, implemented by parents, had on the language development of 2 to 7 year olds. The studies were reviewed for methodological quality and relevance (Gough, 2007), and conclusions and recommendations were made. Evidence reported in this review demonstrates that DR with children, implemented by parents, can be endorsed as good practice (Gersten et al., 2005).

2.0 Introduction

2.1 Role of parents

For the majority of young people, the family environment is the very first context in which they can begin to develop learning behaviours (Sammons et al., 2015). Positive engagement by parents and/or carers in their child's early learning goals and the quality of the home learning environment have been widely reported as resulting in significantly improved academic outcomes for young people of all school ages (Castro et al., 2015).

A longitudinal study which aimed to investigate the impact that the home learning environment had on children's language development reported that a positive, supporting home environment had a significant influence on the child's educational outcomes, as well as on predictions of later academic attainment (Sammons et al., 2002).

Early child developmental theorists such as Vygotsky have sought to understand how early learning happens, which is beneficial for understanding the elements of the home learning environment that can promote a child's learning. Vygotsky (1978) hypothesised that early childhood learning results from appropriate levels of scaffolding from a person more knowledgeable than the learner, to help move the learner's development forwards. In the case of a child developing typically, the more knowledgeable person who mediates the process of learning is typically the young person's parent or carer.

2.2 Dialogic Reading

The term Dialogic Reading (DR) relates to interactive book-sharing strategies which aims to promote the acquisition of skills for reading such as receptive and expressive language (Vally, 2012) and new word acquisition by scaffolding the development of novel vocabulary (Chow et al., 2008). Adults can support younger learners to develop their language-related skills by engaging in conversations about the text they have read and the themes described within them. DR refers to experienced others (usually 'adults') use of interactive questioning and commenting behaviours whilst sharing books with their child. DR strategies can include tracking the child's interest, initiating conversation through open-ended questions and repeating back key themes to the child (Vally et al., 2015).

DR is made up of four key techniques, utilised by the adult when book-sharing with a child, to push the child to think more widely and in greater depth about the text they are reading. The acronym "PEER" refers to these techniques: "Prompting by the adult; Evaluating what the child said; Expanding on the child's response; and Repeating back or asking the child to repeat back" (Towson & Gallagher, 2014).

2.3 Rationale and Relevance

Before a child attends any educational setting, they experience a multitude of learning opportunities in their home learning environment. Over time, and with appropriately scaffolded support by a knowledgeable adult (such as parent or carer), children use these early experiences to acquire novel language and develop language-related cognitive skills (Fox et al., 2010).

(Indeed, research has shown that without positive contributions of parents and a positive home learning environment, children are less likely to reach their academic potential, and meet the demands set out for them by living in a demanding society (Castro et al., 2015). As such, parental involvement is often deemed to be a defining variable in which predictions of future academic attainment can be based and therefore, it is imperative that schools and families understand the influence that parental involvement can have on the child's later outcomes. Furthermore, understanding of the significant lasting impact that parents and/or carers can have on the progress of a child's language over time, (language being a powerful predictor of academic success (Graham, 1987)) should be prioritised in order to support parents to support their children's learning to the best of their ability.

There is an abundance of research into the role that DR strategies can have on both the acquisition and development of language skills by the child and the developing relationship in the adult-child dyad. There are two previous reviews investigating the impact of DR on language skills (Pillinger & Vardy, 2022; Towson et al., 2017) and both lacked a significant number of studies utilising control groups and therefore, results may not have been the result of a DR intervention, rather the result of an extraneous variable. Thus, this review ensured that included studies utilised a Randomised Controlled Trial method. Further, the previous reviews contained studies which were implemented in a school class room setting, rather than in the child's home context. By focussing on the classroom context of DR, previous reviews failed to acknowledge the significant role that parents can have in supporting their child's language development.

This review is of particular relevance to Educational Psychologists (EPs) as it can provide evidence for involving parents in delivering DR interventions which aim to promote and advance children's language development. Parents and carers can play a role in their child's early language development (Topping et al., 2013) and by encouraging parents to take an active role in this development, we are not only expediting this progress, but also providing opportunities for secure attachments and social skills to develop through book-sharing (Hoyne & Egan, 2019).

2.4 Research Question

What is the impact of Dialogic Reading intervention between parent and child on the child's language development?

3.0 Critical Review of the Included Studies

3.1 Systematic Literature Search

For this review, a systematic literature search was completed on 10th December 2022 on four electronic databases: Web of Science, ERIC, PsycInfo and EBSCO. Table 1 displays search terms used for the literature search for this review. These terms were used in both subject-heading searching and specific word searching.

Table 1

Search Terms for Literature Search

Intervention		Sample		Sample age
"Parent* child read*"	AND	"Language develop*"	AND	"Child*"
"Home read*"	AND	"Language learn*"	AND	"Primary school-aged"
"Read* at home*"	AND	"Language develop"		
"Parent* child read*"	AND	"Language develop*"	AND	"Child*"
"Home read*"	AND	"Language develop*"		
"Read* at home*"	NEAR/2	"Language difficulty"		
"Dialogic read*"	AND	"Parent"	AND	"Child*"

Eight separate searches were run to yield as many returns as possible, using a number of combinations of search terms and operators. The results of these searches were then combined, and any duplicates were removed. Some search terms were combined (such as "read* at home" was combined with "language develop*" and "language difficulty") to ensure the search returned literature that related to all aspects of the research question, and not just one. For example, if the search had just used "read* at home" and "language difficulty", but without including "language develop*", then the searches would not necessarily relate to the development of language, and therefore would not be relevant to the research question. One search used the operator NEAR/2, to increase the number of returned searches, however this made very little impact (NEAR/2 only added two searches, and they were publications that had already been previously identified), so was not deemed necessary to use this operator again. Sample age was added to some search

terms, to ensure the results related to the research question, which specifically focussed on language development in pre-school and primary school-aged children.

Table 2

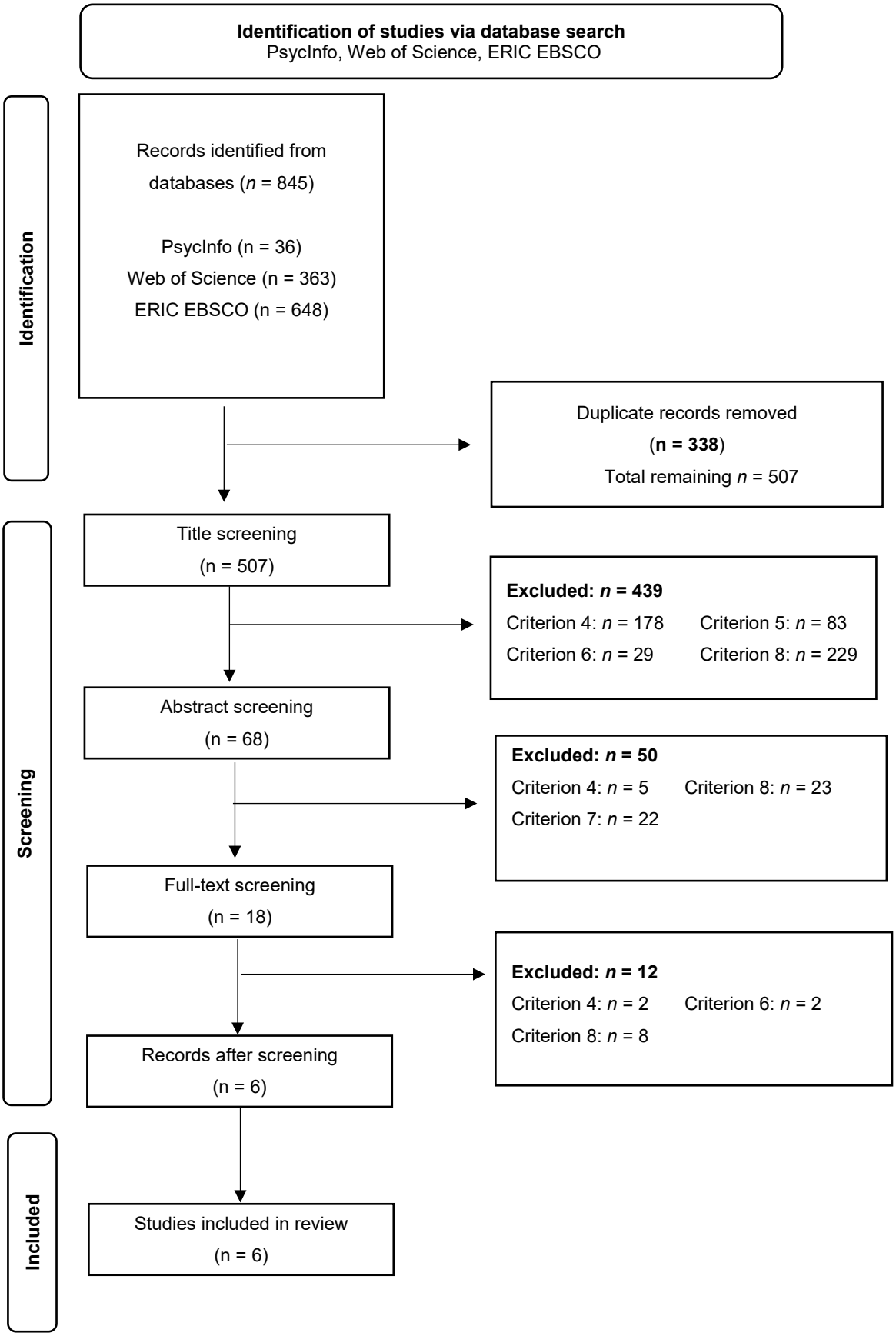
Inclusion and Exclusion Criteria

Inclusion characteristic	Inclusion criteria	Exclusion criteria	Rationale for criteria
1. Publication type	The study must have been published in a journal that has been peer-reviewed.	The study has not been published in a journal that has been peer-reviewed.	To ensure there is high methodological quality.
2. Language	The study is written in the English language and has not been translated.	The study is published in a language other than English.	Ensures that the study is written exactly how the authors intended and translation has not altered any meaning.
3. Date of Publication	The study was published between 2012 and present day.	The study was published prior to 2012.	The aim of this review is to appraise the most recent evidence in this field.
4. Intervention	One of the conditions of the intervention must include Dialogic Reading between parent (or carer) and child.	The study does not include Dialogic Reading between parent (or carer) and child in one of the conditions.	The aim of this review is to appraise the effectiveness of Dialogic Reading in parent-child dyads.
5. Age of participants	Children are pre-school or primary school aged (between 2 and 11 years old).	Children are older than 11 years old.	The aim of this review is to understand the impact that parent-led DR strategies can have on children's language skills.

Inclusion characteristic	Inclusion criteria	Exclusion criteria	Rationale for criteria
6. Setting intervention is delivered in	The Dialogic Reading delivery occurs in the home environment.	The Dialogic Reading delivery occurs in an environment other than the child's home.	The aim of this review is to appraise the evidence of DR interventions practiced in the child's home.
7. Outcome variables measured	The study includes at least one outcome measure relating to the acquisition or development of a skill relating to language.	Outcomes unrelated to language acquisition or development.	The aim of this review is to understand the impact that parent-led DR strategies can have on children's language skills.
8. Study design	The study makes use of a Randomised Controlled Trial (RCT) design.	The study has not used the RCT design.	RCTs have been described as the "gold standard" (Ginsburg & Smith, 2016) for gathering experimental research and the stringent design can help to reduce selector bias (Hariton & Locascio, 2018).
9. Person delivering DR to child	The study uses parents and/or carers to deliver DR to the child.	The study uses anyone other than parents and/or carers to deliver DR to the child.	The review aims to understand the unique impact that parents delivering DR has on child language development.

Figure 1

PRISMA Chart displaying the article Screening Process



Overall, 845 studies were returned from the four databases. After removing duplicates (*n* = 338), 507 articles remained which were then screened at title-level and 439 studies were excluded for not meeting the inclusion criteria. Next, 68 articles were screened at abstract level and 50 studies were excluded as they did not meet the inclusion criteria. At full-text screening, 18 articles were examined and 12 were excluded for not meeting the inclusion criteria, leaving a remaining6 articles after screening. The final review appraised 6 studies. Figure 1 displays an overview of this process. Table 2 shows inclusion and exclusion criteria for this review.

One previous systematic literature review was found that had already explored this topic (Pillinger & Vardy, 2022). However, this review included studies which did not include control groups (only 23 out of the 44 studies reviewed, reported the use of a control group) (Pillinger & Vardy, 2022). The current review therefore extends previous literature by ensuring the studies included a control group to which the participants are randomly assigned. This is important to ensure that we can determine that the results of DR with children are the result of the DR and not another factor.

3.2 Included Studies

Following the screening process, six studies remained and have been included in this review. See Table 3 for references of studies that were included in this review. See Appendix A for references of studies excluded at full-text screening (*n* = 12).

Table 3

References for Final Studies Included in This Review

	Reference
1.	Brannon, D. & Dauksas, L. (2014). The Effectiveness of Dialogic Reading in Increasing English Language Learning Preschool Children's Expressive Language. *Research in Early Childhood Education*, 5(1), 1 – 10.
2.	Chacko, A., Fabiano, G., Doctoroff, G. & Fortson, B. (2018). Engaging Fathers in Effective Parenting for Preschool Children Using Shared Book Reading: A Randomized Controlled Trial. *Journal of Clinical Child and Adolescent Psychiatry*, 47(1), 97 – 93.
3.	Kotaman, H. (2020). Impacts of Dialogical Storybook Reading on Young Children's Reading Attitudes and Vocabulary Development. *Reading Improvement, 57(1),* 40 – 45.
4.	Sim, S., Berthelsen, D., Walker, S., Nicholson, J. & Fielding-Barnsley, R. (2013). A shared reading intervention with parents to enhance young children's early literacy skills. *Early Child Development and Care, 184(11),* 1531 – 1549.
5.	Towson, J. & Gallagher, P. (2014). Training Head Start parents in dialogic reading to improve outcomes for children. *International Journal for Child Health and Human Development*, 7(3), 287 – 296.

Reference
6. Vally, Z., Murray, L., Tomlinson, M. & Cooper, P. (2015). The impact of dialogic book-sharing training on infant language and attention: a randomized controlled trial in a deprived South African community. *Journal of Child Psychology and Psychiatry, 6(3),* 865 – 873.

3.3 Mapping the Field

Studies included in this review aimed to investigate the effectiveness of DR implemented by parents in supporting children's language acquisition and development. The included studies varied somewhat in their population sample, intervention procedure and outcome measures. Key features from each study can be found in Table 4.

Table 4

Mapping the Field

Author(s) and date	Location	Sample	Research design	Intervention type and delivery	Key findings
Brannon & Dauksas (2014)	A school in the Midwest United States	**Size:** 40 parents, 43 preschool children **Gender:** Children: 26 M, 17 F **Age:** Children: between 3 and 5 years old. **Ethnicity:** Unknown **Other:** 75% DR group and 61% of control group speak Spanish at home.	RCT. Pre-test and post-test. **Experimental conditions:** DR (n = 22), FTG control (n = 21).	**Type:** 15 minutes DR training every Monday focussing on the DR strategies. Watched the DR method being used in video clips every Tuesday. Every Wednesday new book given. Literacy interactions between parent and child video-recorded. **Delivery:** DR training delivered 3 days a week for 10 weeks.	Children in the DR group acquired more new words in comparison to the control group.

Author(s) and date	Location	Sample	Research design	Intervention type and delivery	Key findings
Chacko et al. (2018)	Three Head Start centres in New York.	**Size:** 126. **Gender:** Children: FSSP = 66% M, Waitlist = 69% M **Age:** FSSP mean age = 4.76; Waitlist mean age = 4.42 Fathers mean age: FSSP = 36.77; Waitlist = 35.25 **Ethnicity:** FSSP = 85% Hispanic; Waitlist = 89% Hispanic	RCT. Pre-test and post-test. **Experimental conditions:** FSSP (n = 64), Waitlist control (n = 62).	**Type:** Father parenting programme. **Delivery:** FSSP is an 8-week, 90-minute programme. Video-recorded examples of father-child reading interactions were viewed and discussed. Child engaged in arts and crafts whilst fathers attended the training. Sessions focussed on specific DR content.	Children in the FSSP group showed improved auditory comprehension and expressive communication in comparison to the control group.
Kotaman (2020)	Bursa, Turkey.	**Size:** 80. **Gender:** Children: DR = 11 M, 9 F; Control = 12 M, 8 F Parents: DR = 4M, 16 F;	RCT. Pre-test and post-test. **Experimental conditions:** DR (n = 22), FTG control (n = 21).	**Type:** DR intervention and control group (no intervention). **Delivery:** 120 minute DR training including role play of DR techniques.	Significant improvements in receptive language skills for children who had experienced DR strategies condition.

Author(s) and date	Location	Sample	Research design	Intervention type and delivery	Key findings
		Control = 3 M, 17 F **Age:** DR mean age = 3.9 years; Control mean age = 3.9			
Sim et al. (2013)	Queensland, Australia.	**Size:** 80. **Gender:** Children: DR = 42 M, 38 F; Parents: 5 M, 68 F **Age:** Children mean age = 5.53 years	RCT. Pre-test and post-test. **Experimental conditions:** DR group DR + PR group and control group.	**Type:** DR intervention group, DR + PR intervention group and control group (no intervention). **Delivery:** Videotapes of strategies shown to the parents, opportunities for role playing and questions.	Those in the DR and DR + PR groups showed significantly improved scores on language outcome measures. The addition of the PR group did not lead to any significant differences.
Towson & Gallagher (2014)	Three Head Start centres in Southeast United States.	**Size:** 24. **Gender:** Children control: DR = 7 M, 5 F; Children DR = 5 M, 8 F **Age:** Control Children mean age = 47 months;	RCT Pre-test and post-test. **Experimental conditions:** DR group and control group.	**Type:** DR intervention group, control group. **Delivery:** 30-minute training on DR techniques for parents, including watching video clips and picking out the techniques used.	Standard scores for receptive language in the DR group increased (not significant).

Author(s) and date	Location	Sample	Research design	Intervention type and delivery	Key findings
		DR children mean age = 46.54 months.			
Vally et al. (2015)	Khayelitsha, Cape Town.	**Size:** 91. **Gender:** Children control: DR = 26 M, 16 F; Children DR = 33 M, 16 F **Age:** Control Children mean age = 15.29 months; DR children mean age = 15.45 months. Parent control mean age = 31.76 years Parent DR mean age = 33.35 years	RCT Pre-test and post-test. **Experimental conditions:** DR group and control group.	**Type:** DR intervention group, control group. **Delivery:** 8, 90-minute sessions, 1 per week. Training delivered in groups of 4-5 parents and children dyads. Provided with weekly supervision throughout. Included role play, question asking and feedback.	Parents who had had the DR training reported significantly more words understood by their child. Parents who had had the DR training reported significantly more words vocalised by their child.

3.4 Gough's Weight of Evidence

This review aimed to appraise the six studies using a Weight of Evidence (WoE) framework (Gough, 2007). Each study was appraised for methodological quality (WoE A), methodological relevance (WoE B) and topic relevance (WoE C).

WoE A was determined using an adapted version of using Gersten et al. (2005). coding protocol due to its relevance for experimental group designs. Explanation of the protocol used and the final coding protocols are found in Appendix C.

The protocols for WoE B and WoE C were created by the reviewer. Total combined scores from WoE A, B and C were then averaged to provide a total WoE rating (WoE D), shown in Table 5. See Appendix B for criteria for each of the WoE ratings as well as the overall WoE ratings.

Table 5

Total Weight of Evidence Ratings for Studies Included in this Review

Study	WoE A	WoE B	WoE C	WoE D
Brannon & Dauksas (2014)	1 (Low)	2.25 (Medium)	2.66 (High)	1.97 (Medium)
Chacko et al. (2018)	1 (Low)	2.25 (Medium)	2.33 (Medium)	1.53 (Medium)
Kotaman (2020)	1 (Low)	2.25 (Medium)	2.00 (Medium)	1.42 (Low)
Sim et al. (2013)	3 (High)	2.75 (High)	2.33 (Medium)	2.69 (High)
Towson & Gallagher. (2014)	1 (Low)	2.25 (Medium)	2.66 (High)	1.64 (Medium)
Vally et al. (2015)	1 (Low)	2.25 (Medium)	2.66 (High)	1.97 (Medium)

Note:

For WoE D ratings, a high score is a rating of greater than 2.5, a medium score is a rating of greater than 1.5 and a low rating is any score 1.5.

3.5 Critical Review of Included Studies

Participants

A total of 484 parent/carers and their children participated across the studies in this review, with participant numbers ranging between 25 (Towson & Gallagher, 2014) and 126 (Chacko et al., 2018). Three studies (Chacko et al., 2018; Sim et al., 2014, Vally et al., 2015) reported attrition rates (a desirable

quality indicator for experimental research designs, Gersten et al., 2005), and particularly useful to understanding the validity of the results. All three studies had attrition rates of less than 30% (a further desirable quality indicator, Gersten et al., 2005). The children in the samples were all aged between 2 and 7 years old. One study (Towson & Gallagher, 2014) included pre-school aged children (mean aged 47.00 months for control group and 46.54 months for experimental group). All studies in this review utilised parents in their sample, as per the review question. One study sample (Chacko et al., 2018) was made up of Fathers only and one study included aunts and neighbours in their parent/carer sample (Vally et al., 2015). Further, one study (Towson & Gallagher, 2014) included both neurotypical children as well as children with disabilities. No other studies reported on the neurodiversity of the child samples.

Samples were recruited in three main ways: through Head Start centres (Chacko et al., 2018; Towson & Gallagher, 2014); via the child's school (Brannon & Daksaus, 2014; Kotaman, 2020; Sim et al., 2013) and by looking at children living in a specific geographical area (Vally et al., 2015).

Setting

All studies included samples recruited from OECD countries, meaning that although the studies were not conducted in the UK, they were conducted in countries with similar demographics and socio-economic status and therefore results are more generalisable to the UK. Three studies were completed in schools in the United States (Brannon & Dauksas, 2014; Chacko et al., 2018;

Towson & Gallagher, 2014) one in Turkey (Kotaman, 2020), one in Australia (Sim et al., 2014) and one in Khayelitsha, Cape town (Vally et al., 2015).

The DR training for all six studies was completed in educational settings (two in Head Start centres (Chacko et al., 2018; Towson & Gallagher, 2014) and the remaining four in primary schools. One study included children attending a private school (Kotaman, 2020). In all six studies, the subsequent DR application was implemented in the child's home over a number of weeks.

Study design

All six of the studies utilised a randomised controlled trial procedure as it is considered the "gold standard design" for experimental research (Ginsburg & Smith, 2016). For this reason, all of the studies were given high WoE B ratings, as RCTs have been widely reported to be the most advantageous design for answering questions involving two or more experimental groups (Ginsburg & Smith, 2016). RCTs also help reduce the possibility of selection and administrator bias (Hariton & Locascio, 2018) which is beneficial to ensure reliable and valid conclusions about the effectiveness of DR interventions for developing child language, can be drawn.

The six studies in this review ascribed the control groups to a number of different activities. Two studies (Kotaman, 2020; Vally et al., 2015) utilised a passive control group whereby the control did not receive any input from the research teams, other than assessments. Two studies used an active control (Brannon & Daksaus, 2014; Towson & Gallagher, 204), whereby the control group followed an alternative intervention, in this case, reading with no dialogic instruction (Brannon & Daksaus, 2014) and positive behavioural or

maths training (Towson & Gallagher, 2014). These studies received higher WoE B scores as they helped to ensure that any effects were the result of dialogic reading and not another factor, such as just receiving attention.

Intervention

As per the inclusion criteria, DR was implemented by parents and/or carers, in all six studies. The DR training for all six studies involved teaching parents about the relevant DR strategies including prompts for the child to continue the reading, prompts for the child to recall previously read text, and wh-questions (Cheng et al., 2021) in order to increase child reading independence.

Only two of the six studies (Chacko et al., 2018; Sim et al., 2014) reported intervention fidelity in their research, giving them higher WoE A ratings for the intervention fidelity subcategory. The reporting of intervention fidelity in published research is imperative to ensure that readers can accurately draw conclusions about the quality, reliability and reproducibility of a study (Murphy & Gutman, 2012). Chacko et al (2018) and Sim et al. (2014) both reported the use of videotaping as a tool to monitor and understand the fidelity of their interventions. Additionally, Chacko et al. (2018) also provided additional training for the parents when the fidelity ratings fell below 80% for two or more intervention sessions.

Intervention delivery

Interventions implemented in the six studies ran for between 5 and 10 weeks. One study (Kotaman, 2020) did not report the duration of the intervention and

therefor received a lower WoE B rating for this category. All studies in this review utilised a pre-test and post-test assessment design, with one study (Sim et al., 2014) also including a follow-up assessment 7 weeks after the post-test assessments, which can help to provide information on intervention efficacy and longer-term benefits (Lewellyn-Bennett et al., 2016). The frequency of the intervention instruction also varied throughout the intervention period amongst the six studies.

Brannon and Dauksas (2014) delivered 5 weeks of instruction over a period of 10 weeks with parents receiving DR strategy training for 15 minutes every Monday and watching modelling of DR for 15 minutes every Tuesday. Chacko et al. (2018) implemented an 8-week, 90 minute per week parenting programme for fathers, which incorporated DR training. Sim et al. (2014) delivered the DR intervention to parents for hour on one occasion. Later, follow-up conversations were had with all the parents to ensure they understood the DR techniques to use when reading with their child. Kotaman (2020) gave DR instruction to the parents during a two-hour session whereas Towson and Gallagher (2014) delivered the DR training to parents in a 30-minute session and Vally et al. (2015) delivered their training in a series of 90-minute sessions, every week for 8 weeks. The last 20 minutes of each session focussed on the book they were going to be reading with their child for that week.

All six studies delivered the intervention in groups of parents and all used videotaped vignettes in order to observe DR in action and provide opportunities to ask relevant questions. All studies provided the parents with

books to read with their child, so no family was expected to source books for themselves. This may have helped reduce attrition rates for families of low socio-economic status who otherwise could struggle to buy new books for their children.

All studies discussed in this review all described the DR intervention in detail and therefore, they all received higher WoE A ratings as the studies had high reproducibility. However, as Sim et al. (2014) study had a follow-up assessment, and therefore the long-term effects of parents' DR training on language acquisition could be measured, which resulted in a higher WoE A rating for this study.

Outcome Measures

The reviewed studies utilised a number of outcome assessments, collecting data on a range of post-intervention measures. As per the inclusion criteria, all studies included at least one measure of a skill relating to the improvement of children's language. Several studies (Chacko et al., 2018; Sim et al., 2013; Kotaman, 2020; Sim et al., 2013; Towson & Gallagher, 2014) used two or more assessments measuring the child's language and therefore received higher WoE B ratings for this sub-category. A range of other outcome variables not relevant to this review were also measured, including parental stress and depression (Chacko et al., 2018) and sustained attention (Vally et al., 2015).

Brannon and Dauksas (2014) used the Individual Growth and Development Indicator (IGDI) to assess the child's picture naming skills. The authors report that this was a reliable measure of language in children, with one-month

alternate for reliability coefficients ranging between *r* = .44 to .78. Further, the IGDI was found to correlate with other norm-referenced language skill measures including the Peabody Picture Vocabulary Test.

Chacko et al. (2018) used the Preschool Language Scales (Fourth Edition) to assess children's language skills. The authors did not report on the reliability and validity of the outcome measures, however, authors of the tool report it to have very good internal consistency, (α = >.90).

Kotaman (2020) and Sim et al. (2014) and Towson and Gallagher (2014) used the Peabody Picture Vocabulary Test (PPVT) to assess children's language skills. Kotaman (2020) and Towson and Gallagher (2014) did not report reliability or validity statistics for this assessment tool, however, Sim et al. (2014) reported test-retest reliabilities of .91 and .92 for forms A and B respectively.

Sim et al. (2014) also used the Hundred Picture Naming Test (HPNT) to assess children's expressive language and reported a test correlation with expressive vocabulary is .83. Further, Sim et al. (2014) used the Phonological Abilities Test (PAT) to assess the child's phonological skills, and reported test-retest reliability scores of .58 - .86. Finally, the authors also used the Concepts About Print (CAP) to assess children's print awareness skills, and reported test-retest reliability between .73 and .89 for the different scales.

Towson and Gallagher (2014) used the Expressive One-Word Picture Vocabulary Test – Fourth Edition (EOWPVT) to assess children's expressive vocabulary skills. The authors did not report any reliability or validity

statistics, however, the manual reports reliability coefficients ranging between .93 and .98.

Vally et al. (2015) used the Communicative Development Inventory to interview parents about their children's language skills. The authors did not report a reliability measure for this tool, however, previous testing has reported good test-retest reliability of .86 - .95 (Dale et al., 1989). Vally et al. (2015) also developed an assessment of language comprehension skills, based on the PPVT. The authors did not report reliability coefficients for this measure and these cannot be sourced elsewhere due to the measure being original.

Findings and Effect Sizes

Table 7 displays a summary of outcome measures, descriptions of findings, effect sizes and their corresponding descriptors. Outcome measures which did not directly relate to the aims of this review (child language outcomes) can be seen in Appendix C.

One study (Brannon & Dauksas, 2014) did not report an effect size and this had to be calculated by the reviewer using the data reported. Two studies reported (Chacko et al., 2018 and Sim et al., 2014) reported Cohen's *d.* Two studies (Towson & Gallagher, 2014 and Vally et al., 2015) reported Partial Eta Squared, which the reviewer converted to Cohen's *d*.

All six of the studies reported at least one outcome measure relating to the child's language development after receiving DR intervention from their parent or carer. One study (Towson & Gallagher, 2014) did not find a

significant relationship between DR intervention and child's language development, but found a significant relationship between DR intervention and other outcome measures unrelated to this review. Of the studies that reported significant effects of parent-implemented DR intervention, one reported only small effect sizes for the outcome measures, suggesting there may be limited practical applications of the intervention in the real world.

Effect sizes were calculated (where not reported) for all six studies in the review using the standardised mean difference (Cohen's *d*). The effect size was a measure of the strength relationship between DR interventions and the positive increase in language-related skills in young children. For the six studies, the effect size relates to the post-DR intervention measures for participants in both the experimental and control group. For studies that did not report Cohen's *d*, it was calculated using Psychometrica (Lenhard & Lenhard, 2017). Descriptors for the effect sizes are provided in Table 6.

Table 6

Standardised mean difference (Cohen's *d*) descriptors (Cohen, 1992)

Effect Size	Descriptor
.2	Small
.5	Medium
.8	Low

Table 7

Summary of Outcomes, Main Findings, Effect Size (Cohen's d) and Descriptors for included studies

Study	Outcome measures	Main findings	Effect Size (Cohen's *d*) *	Effect Size Description	WoE D
Brannon et al. (2014)	Child's attention to text (Adult – Child Interactive Reading Inventory)	Dialogic Reading training for parents had a significant positive influence on their reading interaction style with their child. Dialogic Reading training with parents had a positive influence on their child's use of expressive language at post-test, in comparison to parents who had not been trained in Dialogic Reading.	.43	Small	1.97 (Medium)
	Increase in child's expressive language (Picture-naming subtest of the Individual Growth Developmental Indicators)		2.56	High	
			1.17	High	
			.32	Small	
Chacko et al. (2018)		Parenting programmes involving Dialogic Training for fathers had a significant impact on father nurture, positive parenting, child behaviour, child auditory language and child's expressive language.	.12	Small	1.53 (Medium)
	Child auditory language (measured using the Preschool Language Scales) *		*.82*	Large	
	Child use of expressive language (measured using the Preschool Language Scales) *		.91	Large	

Study	Outcome measures	Main findings	Effect Size (Cohen's *d*) *	Effect Size Description	WoE D
		*Significantly different from the mean of the control group at post-test (p<.01).	.04	Small	
			.10	Small	
			.63	Medium	
			.53	Medium	
			.34	Small	
			.51	Medium	
			.52	Medium	
			.51	Medium	

Study	Outcome measures	Main findings	Effect Size (Cohen's *d*) *	Effect Size Description	WoE D
Kotaman (2020)	Child's receptive vocabulary (measured using a Turkish version of the Peabody Picture Vocabulary Test)	Children whose parents had been trained in Dialogic Reading showed a significant increase in receptive vocabulary in comparison to those whose parents were not Dialogic Reading trained.	.75	Medium	1.42 (Low)
	Child's reading attitudes (measured using the Preschool Reading Attitudes Scale)		.40	Small	
Sim et al. (2013)	Child receptive language (Peabody Picture Vocabulary Test)	Dialogic training for parents improved children's language skills. These results were maintained 3 months later at follow-up. There was no significant difference of the addition of print referencing to DR training on the development of children's language skills.	.35	Medium	2.69 (High)
	Child expressive language (Hundred Pictures Naming Test)		.21	Small	
	Child's phonological skills (measured using the Phonological Abilities Test)		.28	Small	
	Print awareness skills (measured using the Concepts About Print)		.23	Small	

Study	Outcome measures	Main findings	Effect Size (Cohen's *d*) *	Effect Size Description	WoE D
Towson & Gallagher (2014)	Child receptive language (Picture Vocabulary Test in English) and the Test de Vocabulario en Imagenes Peabody in Spanish) Child expressive vocabulary (Expressive One-Word Picture Vocabulary Test and the Expressive One-Word Picture Vocabulary Test – Spanish-Bilingual Edition) Preliteracy skills (measured using the Get Ready to Read! – Revised)	There were no significant differences between children whose parents had received DR training and the control group, for tests of receptive language. However, the standard scores on these assessments for the experimental group did increase, whereas the scores for the control group remained relatively stable. Similarly, for tests of expressive language, small improvements were made for the experimental group but not for the control group. These changes were not significant.	.27 1.19 .57	Small Large Medium	1.64 (Medium)

Study	Outcome measures	Main findings	Effect Size (Cohen's *d*) *	Effect Size Description	WoE D
Vally et al. (2015)	Child language skills (measured by interviewing parent using the Mac-Arthur Bates Communication Development Inventory)	Dialogic reading between parent and child had a moderate impact on child lexical production and a large impact on child comprehension, for children whose parents had received DR training. There was no significant difference between the reported change in receptive and expressive language for children whose parents had received DR training.	*1.09*	Large	1.53 (Medium)
	Child receptive vocabulary (measured using author's own assessment, based on the Peabody Picture Vocabulary Test Revised)		.41	Medium	
			1.19	Large	

*Cohen (1992) states that the effect size is 'Small if <0.3, 'Medium' if 0.31 – 0.5 and 'Large if >0.5.

4.0 Conclusions and Recommendations

4.1 Discussion of Findings

This review aimed to evaluate whether Dialogic Reading, led by parents (and/or carers) had an effect on their child's language-related skills in preschool and early school-aged children. Six studies met the inclusion criteria. Studies were appraised on their methodological quality (WoE A), their methodological relevance (WoE B), their relevance of the study to the review question (WoE C) and these were averaged to provide an overall score for WoE D (Gough, 2007).

The two studies that received the largest WoE D ratings (Brannon & Dauksas, 2014 and Sim et al., 2013) reported mixed effect sizes. Brannon and Dauksas (2014) reported high effect sizes for parents use of literacy strategies (d= 1.17) and small effect size (d= 0.32) for increase in child's expressive language whereas Sim et al. (2013) reported a medium effect size for child's use of receptive vocabulary (d= 0.35), and a small effect size for child's expressive vocabulary (d=0.21). This suggests that despite the studies being the most methodologically sound and relevant to this review, the strength of the relationship between DR and child's use of receptive language (d=0.35, Sim et al., 2013), may be weak. However, one study (Vally et al., 2015) reported a large effect size for child language skills following DR (d= 1.09) and had a medium WoE D, suggesting that the higher methodological quality and relevance may have had an effect on the strength of the relationship between DR intervention and early language development.

4.2 Recommendations for Practice

Gomes-Koban et al., 2019 reported a need to create a sustainable way for researchers and the wider educational settings and communities to collaborate, in order for empirical research to be conducted and the findings utilised to support evidence-based interventions in the classroom. Therefore, when considering whether the pertinence of dialogic reading interventions with parents, it is important to consider whether results of empirical data can be generalised to wider settings. The six studies all demonstrated strengths within the areas of methodological quality and relevance, which resulted in medium or high WoE for all studies. Therefore, there is initial promising research to tentatively promote the use of DR with parents to support child language development.

One study (Chacko et al., 2018) utilised a parent sample made up of fathers only. This novel contribution to the DR literature gave evidence to support the inclusion of fathers in research that is otherwise dominated by female parent/carer samples. Research has shown the role of fathers is most significant for low-income families (Tamis-LaMonda et al., 2013) and as EPs play a role in supporting at risk groups, this is may be of particular interest to the profession. Therefore, the applicability of this intervention to these groups could be important within their practice, and particularly to support the most vulnerable groups.

4.3 Limitations of the Review

In this review, only one study included follow-up assessment at a later time point (essential criteria for WoE A, Gersten et al. (2005)). With this in mind, it is difficult to understand the long-term benefits that DR with parents can have on a child's language development. However, the study that included a later follow-up reported that the benefits in child language development reported immediately post-intervention, could also be seen at follow-up three months later. Therefore, tentative conclusions can be made so suggest that DR with parents can have longer-term positive impacts on the child's language development. Nonetheless, if the studies in this review had included follow-up assessments, they would have provided more information on the longer-term impacts of DR on language development and would have improved the WoE A and overall WoE D ratings of the study.

The geographical location of the studies could be seen as a limitation of this review. None of the studies were conducted in the UK, however, all of the studies were conducted in OECD countries to ensure likeness to the UK. In order for results to be generalisable to educational psychology practice in the UK, it would be beneficial to use samples collected from the UK.

Finally, the development of a WoE C protocol to assess the studies' application to the research question, could be challenged. To date, there is no published protocol to assess WoE C and therefore, it was at the discretion of the reviewer to create this. With this in mind, the protocol has not been tested for construct validity. However, the relevant constructs known to affect

study relevance (including setting, sample and outcome measures) were included (Newman & Gough, 2020).

4.4 Recommendations for Future Research

In order for research to be generalised and applicable to educational psychology practice in the UK, further research is required utilising samples from the UK in order to provide a suitable evidence-base for the use of parent-led DR interventions. Whilst this review did not include studies conducted in the UK, it did include studies conducted in OECD countries, known to have similar educational systems to that of the UK.

Research by Chacko et al. (2018) showed the importance of encouraging the commitment of fathers or male carers in research relating to child development. Historically, parent-based samples have been mainly mothers, thus diminishing the significance that the role of fathers can have on their child's growth. By encouraging father-based samples, not only will research on father-influence expand, but we will gain a greater understanding on the differences between mother and father parenting strategies and the impact that these can have on their child's language development.

Future studies should also endeavour to include follow-up measures to evaluate the maintenance of the benefits of DR on child language development. This is particularly important to ensure that interventions recommended by educational psychologists have long-term benefits and face validity.

5.1 References

American Psychological Association. (2005). *APA Policy Statement on Evidence-Based Practice in Psychology.* Retrieved from: http://www.apa.org/practice/guidelines/evidence-basd-statement.aspx

Borenstein, M., Hedges, L., Higgins, J. & Rothstein, H. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1(2),* 97 – 111. Doi: 10.1002/jrsm.12

Brannon, D. & Dauksas, L. (2014). The Effectiveness of Dialogic Reading in Increasing English Language Learning Preschool Children's Expressive Language. *Research in Early Childhood Education, 5(1),* 1 – 10.

Castro, M., Exposito-Casas, E., Lopez-Martin, E., Liasoain, L., Navarro-Asencio, E. & Gaviria, J. (2015). Parental Involvement on Student Academic Achievement: A Meta-Analysis. *Educational Research Review, 14(1),* pp. 33 – 46. Doi: 10.106/j.edurev.2015.01.002

Cohen, J. (1992). A power primer. *Psychological bulletin, 112(1),* 155. Doi: 10.1037//0033-2909.112.1.1.155

Chacko, A., Fabiano, G., Doctoroff, G. & Fortson, B. (2018). Engaging Fathers in Effective Parenting for Preschool Children Using Shared Book Reading: A Randomised Controlled Trial. *Journal of Clinical Child & Adolescent Psychology, 47(1),* 79 – 93. Doi: 10.1080/15374416.2016.1266648

Cheng, L., Xu, W., Gao, Q., Ma, X. & Zhang, Y. (2021). Effects of Dialogic Reading on the Creativity Development of a Chinese Student. *Creative Education, 12(10),* 2371 – 2389. Doi: 10.4236/ce.2021.1210179

Chow, B., McBride-Chang, C., Cheung, H. & Chow, C. (2008). Dialogic reading and morphology training in Chinese children: Effects on language and literacy. *Developmental Psychology, 44(1),*. 233 – 244. Doi: 10.1037/0012-1649.44.1.233

Dale, P., Bates, E., Reznick, J. & Morisset, C. (1989). The validity of a parent report instrument of child language at 20 months. *Journal of Child Language, 16(1),* 239 – 249. Doi: 10.1017/s0305000900008680

Fox, S., Levitt, P. & Nelson, C. (2010). How the timing and quality of early experiences influence the development of brain architecture. *Child Development, 81(1),* 28 – 40. Doi: 10.1111/j.1467-8624.2009.01380.x

Gazzillo, F., Schimmenti, A., Formica, I., Simonelli, A. & Salvatore, S. (2017). Effectiveness is the gold standard of clinical research. *Research in Psychotherapy, 20(2),* 278. Doi: 10.4081/ripppo.2017.278

Gersten, R., Compton, L., Compton, D., Coyne, M., Greenwood, C. & Innocenti, M. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional Children, 71(2),* 149 – 164. Doi: 10.1177/001440290507100202

Ginsburg, A. & Smith, M. (2016). Do randomised controlled trials meet the "gold standard". *American Enterprise Institute.* Retrieved: 06 Janaury,

2023 from: https://www.carnegiefoundation.org/wp-content/uploads/2016/03/Do-randomized-controlled-trials-meet-the-gold-standard.pdf

Gomes-Koban, C., Calet, N. & Defior, S. (2019). Intervention programs in Educational Psychology: Bridging research and practice. *Annals of Psychology, 35(3),* 378 – 388. Doi: 10.6018/analesps.35.3.327941

Graham, J. (1987). English Language Proficiency and the Prediction of Academic Success. *TESOL Quarterly, 21(3),* 505 – 521. Doi: 10.2307/3586500

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. Doi: 10.1080/02671520701296189

Hariton, E. & Locascio, J. (2016). Randomised controlled trials – the gold standard for effectiveness research. *BJOG: an international journal of obstetrics and gynaecology, 125(13),* 1716. Doi: 10.1111/1471-0528.15199

Hoyne, C. & Egan, S. Shared Book Reading in Early Childhood: A Review of Influential Factors and Developmental Benefits. *An Leanbh Og, 12(1),* 77 – 92.

Kotaman, H. (2020). Impacts of Dialogical Storybook Reading on Young Children's Reading Attitudes and Vocabulary Development. *Reading Improvement, 57(1),* 40 – 45.

Kunz, R., Vist, G. & Oxman, A. (2007). Randomisation to protect against selection bias in healthcare trials. *Cochrane database of systematic reviews, 2(1).* Doi: 10.1002/14651858.MR000012.pub2

Llewellyn-Bennett, R., Bowman, L. & Bulbulia, R. (2016). Post-trial follow-up methodology in large randomised controlled trials: a systematic review protocol. *Systematic Reviews, 5(1),* 214. Doi: 10.1186/s13063-018-2653-0

Lopez-Lopez, J., Page, M., Lipsey, M. & Higgins, J. (2018). Dealing with effect size in multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods, 9(3),* 336 – 251. Doi: 10.1002/jrsm.1310

Miller, D., Thomas, M., Waller, M., Nepomnyaschy, L. & Emory, A. (2020). Father Involvement and Socioeconomic Disparities in Child Academic Outcomes. *Journal of Marriage and Family, 82(2),* 515 – 533. Doi: 10.1111.jomf.12666

Murphy, S. & Gutman, S. (2012). Intervention Fidelity: A Necessary Aspect of Intervention Effectiveness Studies. *American Journal of Occupational Therapy, 66(4),* 387 – 388. Doi: 10.5014/ajot.2010.005405

Newman, M. & Gough, D. (2020). Systematic reviews in educational research: Methodology, perspectives and application. *Systematic reviews in educational research,* 3 – 22. Doi: 10.1007/978-3-658-27602-7_1

Nunes, D., Whalon, K., Jackson, E., Intepe-Tingir, S. & Garris, S. (2022). Effects of Dialogic Reading on the Vocabulary Knowledge of

Preschoolers with Autism Spectrum Disorder. *International Journal of Disability, Development and Education*, *69(1),* 282 – 301. Doi: 10.1080/1034912X.2021.1986808

Panter-Brick, C., Burgess, A., Eggermann, M., McAllister, F., Pruett, K. & Leckman, J. (2014). Practitioner Review: Engaging Fathers Recommendations for a game change in parenting interventions based on a systematic review of the global evidence. *Journal of Child Psychology and Psychiatry, 55(11),* pp. 1187 – 1212. Doi: 10.1111/jepp.2014.55.issue-11

Petticrew, M. & Roberts, H. (2003). Evidence, hierarchies and typologies: horses for courses. *Journal of Epidemiology and Community Health, 57(1),* 527 – 529. Doi: 10.1136/jech.57.7.527

Pillinger, C. & Vardy, E. (2022). The Story So Far: A Systematic Review of the Dialogic Reading Literature. *Journal of Research in Reading, 45(4),* pp. 533 – 548. Doi: 10.1111/1467-9817.12407

Pruett, M., Pruett, K., Cowan, C. & Cowan, P. (2017). Enhancing Father Involvement in Low-Income Families: A Couples Group Approach to Preventative Intervention. *Child Development, 88(2),* 398 – 407. Doi: 10.1111/cdev.12744

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Elliot, K. (2002). *Effective provision of pre-school education project (EPPE): Technical paper 8a – measuring the impact of pre-school*

on children's cognitive progress over the pre-school period. Research Report DfES.

Sammons, P., Toth, K., Sylva, K., Melhuish, E., Siraj, I. & Taggart, B. (2015). The Long-Term Role of the Home Learning Environment in Shaping Students' Academic Attainment in Secondary School. *Journal of Children's Services, 19(3),* 189 – 201. Doi: 10.1108/jcs-02-2015-0007

Sim, S., Berthelsen, D., Walker, S., Nicholson, J. & Fielding-Barnsley, R. (2014). A shared reading intervention with parents to enhance young children's early literacy skills. *Early Development and Care, 184(11),* 1531 – 1549. Doi: 10.1080/03004430.2013.862532

Suurmond, R., van Rhee, H. & Hak, T. (2017). Introduction, comparison, and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods, 8(4),* 537 – 553. Doi: 10.1002/jrsm.1260

Tamis-LeMonda, C., Baumwell, L. & Cabrera, N. (2013). *Fathers' role in children's language development.* In: *Handbook of Father Involvement,* pp. 135 – 150. Routledge. Doi: 10.4324/9780203101414

Topping, K., Dekhinet, R. & Zeedyk, S. (2013). Parent-infant interaction and children's language development. *International Journal of Experimental Educational Psychology*, *33(4),* 391 – 426. Doi: 10.1080/01443410.2012.744159

Towson, J. & Gallagher, P. (2014). Training Head Start parents in Dialogic Reading to improve outcomes for Children. *Journal of Child Health and Human Development*, 7(3), pp. 287 – 296.

Towson, J., Jacqueline, A., Fettig, A, Fleury, V. & Abarca, D. (2017). Dialogic reading in early childhood settings: A summary of the evidence base. *Topics in Early Childhood Special Education*, 37(3), pp. 132 – 146. Doi: 10.1177/0271121417724875

Vally, Z. (2012). Dialogic Reading and Child Language Growth – Combating Developmental Risk in South Africa. *South African Journal of Psychiatry, 42(4),* pp. 617 0 627. Doi: 10.1177/00812463120400415

Vally, Z., Murray, L., Tomlinson, M. & Cooper, P. (2015). The impact of dialogic book-sharing training on infant language and attention: a randomised controlled trial in a deprived South African community. *Journal of Child Psychology and Psychiatry, 56(8),* pp. 865 – 873. Doi: 10.1111/jepp.1232

Vygotsky, L. (1978). Interaction between learning and development. In: *Mind in Society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press, pp. 79 – 92.

Wilson, D. (n.d.). *Campbell Collaboration Effect Size Calculator.* Retrieved January 03, 2023, from:http://www.campbellcollaboration.org/esalc/html/EffectSizeCalculator-SMD3.php6.0 Appendices

Appendix A – Excluded studies

Table A1

Table of studies excluded at full-text screening and rationale.

	Excluded study reference	Rationale
1	Cohen, L., Kramer-Via, L. & Frye, N. (2012). Implementing Dialogic Reading with Culturally, Linguistically Diverse Preschool Children. *Research-to-Practice Journal for the Early Childhood Field,* 15(1), 135 – 141.	Criterion 8: Dialogic Reading was implemented by teachers, not by parents as per the review question.
2	Forssman, L. & Gottwald, J. (2022). The impact of interactive book sharing on child cognitive and socio-cognitive development (the REaL trial): study protocol for a randomised controlled trial. *Trials,* 23(1), 802.	Criterion 4: The children were not pre-school aged (10 months old).
3	Asrifan, A., Setiawan, I., Ping, M., Syamdianita, S. & Nurchalis, N. (2022). Dialogic Reading to Promote the Underprivilieged Pre-School Children's Expressive Language Ability. *Script Journal: Journal of Linguistics and English Teaching,* 7(2), 380 – 397.	Criterion 8: Dialogic Reading was implemented by teachers, not by parents as per the review question.
4	Barak, M. & Lefstein, A. (2022). Opening texts for discussion: Developing dialogic reading stances. *Reading Research Quarterly,* 57(2), 449 – 468.	Criterion 8: Dialogic Reading was implemented by teachers, not by parents as per the review question.
5	Simsek, Z. & Erdogan, N. (2015). Effects of the dialogic and traditional reading techniques on children's language development. *Procedia-Social and Behavioural Sciences,* 197, 754 – 758.	Criterion 8: Dialogic Reading was implemented by the researcher, not by parents as per the review question.
6	Urbani, J. (2020) Dialogic reading: Implementing an evidence-based practice in complex classrooms. *Teaching Exceptional Children,* 52(6), 392 – 402.	Criterion 8: Dialogic Reading was implemented by teachers, not by parents as per the review question.
7	Grygas, C., Floyd, K. & Rahn, J. (2018). Dialogic reading and adapted dialogic reading with pre-schoolers with autism	Criterion 8: Dialogic Reading was implemented by the

	Excluded study reference	Rationale
	spectrum disorder. *Journal of Early Intervention,* 40(1), 363 – 379.	researcher, not by parents as per the review question
8	Ganotice Jr, F. A., Downing, K., Mak, T., Chan, B., & Lee, W. Y. (2017). Enhancing parent-child relationship through dialogic reading. *Educational Studies,* 43(1), 51-66.	Criterion 6: This study investigated the impact of dialogic reading on parent-child relationships, not on child language development.
9	Rahn, N. L., Coogle, C. G., & Storie, S. (2016). Preschool children's use of thematic vocabulary during dialogic reading and activity-based intervention. *The Journal of Special Education,* 50(2), 98-108.	Criterion 8: Dialogic Reading was implemented by teachers, not by parents as per the review question.
10	Coogle, C. G., Parsons, A. W., La Croix, L., & Ottley, J. R. (2020). A comparison of dialogic reading, modeling, and dialogic reading plus modeling. *Infants & Young Children,* 33(2), 119-131.	Criterion 8: Dialogic Reading was implemented by teachers, not by parents as per the review question.
11	Chang, C., Hsieh, F., Chen, T., Wu, S., Tzeng, O & Wang, S. (2022). Revisiting Dialogic Reading Strategies with 12-month-old infants. *Early Childhood Education Journal,* 1 – 14.	Criterion 4: The children with whom dialogic reading was completed with, were not of school aged, they were 12 months old.
12	Ping, M. (2014). Group interactions in dialogic book reading activities as a language learning context in preschool. *Learning, Culture and Social Interaction,* 3(2), 146 – 158. Doi: 10.1016/j.lcsi.2014.03.001	Criterion 8: Dialogic Reading was implemented by teachers, not by parents as per the review question.

Appendix B – Criteria and rationale for Weight of Evidence (WoE) Ratings

WoE A: Methodological Quality

WoE A was calculated using Gersten et al's. (2005) coding protocol as it is particularly useful for group experimental research designs. This protocol presents indicators for experimental studies to critically appraise aspects of research articles. This protocol explores key features of the study's methodology including the sample, comparison condition outcome measures and data analysis

Each study were assessed for the 'essential' and 'desirable' criteria outlined in Gersten et al's. (2005) coding protocol. See Table B1 for descriptions of how the WoE A ratings were assigned to each study.

A summary of these scores can be seen in Table B2. A completed coding protocol can be seen in Appendix D.

Table B1

Rating Criteria for WoE A

Rating for WoE A	Criteria outlined by Gersten et al. (2005)
3 (High)	The study meets at least 9 of the essential criteria *and* at least 5 of the desirable criteria.
2 (Medium)	The study meets at least 9 of the essential criteria *and* at least 1 of the desirable criteria.
1 (Low)	The study meets at least 6 of the essential criteria.

Table B2

Overall WoE A Ratings for Studies Included in this Review based on Gersten et al., (2005) protocol

Study	Essential Criteria	Desirable Criteria	WoE A Rating
Brannon & Dauksas (2014)	8	8	1 (Low)
Chacko et al. (2018)	6	5	1 (Low)
Kotaman (2020)	5	2	1 (Low)
Sim et al. (2013)	9	6	3 (High)
Towson & Gallagher (2014)	4	1	1 (Low)
Vally et al. (2015)	7	2	1 (Low)

Note: WoE ratings are categorised as 'High' (3), 'Medium' (2) or 'Low' (1).

WoE B: Methodological Relevance

WoE B aims to appraise the studies' relevance for answering the research question. For this review, WoE B sought to appraise the relevance of the methodology for evaluating the effectiveness of dialogic reading between parents and child on child language development in preschool and early school-aged children.

The protocol used to rate WoE B for each study was developed by the reviewer and is provided in Table 10 below. The reviewer created 5 key criteria: research design, comparison group, outcome measures, intervention detail and follow-up. Scores were averaged across each category. A summary of the scores from the WoE B protocol for each study can be seen in Table B4.

'Research design' considered the design with which the intervention and control groups were delivered, and whether the participants were allocated to an experimental group randomly. Randomisation is a key criterion to reduce the chance of selection bias (Kunz et al., 2007)).

'Comparison group' refers to the extent to which the intervention instruction (dialogic reading training for parents) was isolated in each of the studies included in the review. Hawthorne effects may occur when participants are aware they are being observed and therefore change their behaviour in response (Sedgwick & Greenwood, 2015). Therefore, if the control group received an intervention other than dialogic reading (such as a maths intervention), the Hawthorne effect could be considered to be similar across trials.

'Outcome measures' refers to the extent to which pre and post-intervention assessments provided evidence towards the research question. For this review, the outcome measures refer to the measures relating to the assessment of child language pre and post dialogic reading intervention.

'Intervention details' refers to the clarity of the description of the methodology described in each of the studies. A low score in this category suggests that intervention and control instructions were not clearly described and therefore the ability to replicate the study would be questioned.

'Follow-up' refers to the use of a follow-up assessment following the post-intervention assessment. This is particularly important for researchers to understand the long-term impact of dialogic reading training on children's language development. In order for educational psychologists to recommend the use of dialogic training as a suitable intervention to support child language development, it is important to establish that results are long-lasting.

Table B3

Criteria for WoE B Rating

Criteria	Low - 1	Medium - 2	High - 3
Research design	Single group design with only one group.	Allocation to groups is non-random.	Random allocations of participants to two or more experimental groups.
Comparison group	No comparison or control group.	Control group does not complete any intervention or is a waitlist control group.	Active control group (receiving an alternative intervention).
Intervention details	Intervention and control procedure were not clearly described	Intervention procedure is described in detail but the control group procedure is not clearly described.	Intervention procedure and comparison procedure described in detail.
Follow-up	There is no follow-up assessment following post-intervention	There is a follow-up assessment less than 6 weeks after post-intervention assessment.	There is a follow-up assessment more than 6 weeks after post-intervention assessment

Table B4

Overall WoE B Ratings for Included Studies

Study	Research design	Comparison group	Intervention details	Follow-up	WoE B
Brannon and Dauksas (2014)	3	3	2	1	2.25 (Medium)
Chacko et al. (2018)	3	.3	2	1	2.25 (Medium)
Kotaman (2020)	3	3	2	1	2.25 (Medium)
Sim et al. (2013)	3	3	2	3	2.75 (High)
Towson and Gallagher (2014)	3	3	2	1	2.25 (Medium)
Vally et al. (2015)	3	3	2	1	2.25 (Medium)

Note: An average was taken of the three scores to total the WoE B Rating. A score of < 1.5 is considered 'low', 1.5 – 2.5 is considered 'medium' and a score of > 2.5 are considered 'high'.

WoE C: Topic Relevance

According to Gough (2007), WoE C assesses how relevant a topic is to the review question. The protocol used to code WoE C was developed by the reviewer and can be seen in Table B5. Studies included in this review were assessed on three key areas, including intervention type, language related outcome measures and setting generalisability. These three criteria were deemed by the reviewer to being important to answer the research question. These scores were then averaged to give an overall score for WoE C, shown in Table B6.

Table B5

WoE C Criteria and Rationale

Criteria	Rating	Rationale
Intervention	3 – DR is the only intervention in the experimental group. 2 – DR is combined with another intervention in the main experimental group. 1 – DR is combined with another intervention but is not the key component.	Some studies have combined DR with another intervention, such as PR but this review aims to explore only the contribution of DR on children's language development.
Scope of language related outcome measures	3 – More than two assessments have been used to measure a skill relating to language. 2 – More than one assessment has been used	The research question focuses on the impact of DR on the acquisition and development of children's language skills. Therefore, the outcome measures must be a valid measure of a language skill.

Criteria	Rating	Rationale
	to measure a skill relating to language. 1 – Only one area of language development has been assessed.	
Setting generalisability	3 – The intervention was completed in the UK. 2 – The intervention was completed in another OECD country. 1 – The intervention was not carried out in an OECD country.	So that recommendations can be made to schools within the UK, where the review was written, the study should take place in a country with a similar education to that of the UK. Countries in the OECD are considered to be more similar to the UK, and therefore have similar education systems, in comparison to countries not in the OECD.

Table B6

Overall Ratings for WoE C for studies included in this Review

Study	Intervention	Scope of language related outcome measures	Setting	WoE C
Brannon & Dauksas (2014)	3	3	2	2.66 (High)
Chacko et al. (2018)	2	3	2	2.33 (Medium)
Kotaman. (2020)	3	1	2	2.00 (Medium)
Sim et al. (2013)	2	3	2	2.33 (Medium)
Towson & Gallagher. (2014)	3	3	2	2.66 (High)
Vally et al. (2015)	3	3	2	2.66 (High)

Note: A score of < 1.5 is considered 'low', 1.5 – 2.5 is considered 'medium' and a score of > 2.5 are considered 'high'.

Appendix C – Outcome measures unrelated to the review question

Study	Outcome measure unrelated to child's language
Brannon et al. (2014)	Parents promoting reading (Adult Child Interactive Reading Inventory) Parents use of Literacy strategies when reading with child (Adult – Child Interactive Reading Inventory
Chacko et al. (2018)	Parent expectations (observations using the Parent Behaviour Checklist) Parent discipline (observations using the Parent Behaviour Checklist) Parent nurturing (observations using the Parent Behaviour Checklist) Parental stress (Parenting Stress Index Short Form) Parental depression (Centre for Epidemiologic Studies Depression Scale)

Study	Outcome measure unrelated to child's language
	Positive parenting (Dyadic Parent Child Interaction Coding System)
	Negative parenting (Dyadic Parent Child Interaction Coding System)
	Child problems (Dyadic Parent Child Interaction Coding System)
	Intensity of child behaviour (Eyberg Behaviour Inventory)

Appendix D – Changes made to WoE A Coding Protocol

Changes were made to the WoE A Coding protocol (Gersten et al., 2005) to ensure that all protocol qualities were relevant to the studies included in this review. Omissions from the original protocol can be seen with a strike through the text. This omission was necessary as whether the participants had difficulties or not was not relevant to the review question, nor was it in the inclusion criteria for the review and therefore, this information was not necessary.

Essential and Desirable Quality Indicators for Group Experimental and Quasi-Experimental Research Articles and Reports

Essential Quality Indicators

Participants

1. ~~Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?~~
   - ~~☒Yes (3)~~
   - ~~☐Partly (2)~~
   - ~~☐No (1)~~
   - ~~☐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
   - ☐Unable to Code
3. Was sufficient information provided to characterise the interventionists ~~or teachers~~? Did it indicate whether they were comparable across conditions?

☒Yes

☐No

☐Unable to Code

Essential Quality Indicators

Implementation of the Intervention and Description of Comparison Conditions

4. Was the intervention clearly described and specified?
   ☒Yes

   ☐No

   ☐Unable to Code

5. Was the fidelity of implementation described and assessed?
   ☐Yes

   ☐No

   ☐Unable to Code

6. Was the nature of services provided in comparison conditions described?
   ☒Yes

   ☐No

   ☐Unable to Code

Essential Quality Indicators

Quality Indicators for Outcome Measures

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

   ☐No

   ☐Unable to Code

8. Were outcomes for capturing the intervention's effect measured at the appropriate times?
   ☐Yes

   ☐No

   ☐Unable to Code

Essential Quality Indicators

Data Analysis

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

   ☒Yes

   ☐No

   ☐Unable to Code

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

   ☐Yes

   ☐No

   ☐Unable to Code

Desirable Quality Indicators

1. 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

   ☐Unable to Code

2. 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 (un)familiar to examinees across study conditions?

   ☐Yes

   ☐No

   ☐Unable to Code

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

   ☐Yes

   ☐No

   ☐Unable to Code

4. Was evidence of the criterion-related validity and construct validity of the measures provided?
   ☒Yes
   ☐No
   ☐Unable to Code
5. Did the research team assess not only surface features of fidelity implementation, but also examine quality of implementation?
   ☐Yes
   ☐No
   ☐Unable to Code
6. Was there any documentation of the nature of instruction or series provided in comparison conditions?
   ☒Yes
   ☐No
   ☐Unable to Code
7. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention/?
   ☐Yes
   ☐No
   ☐Unable to Code
8. Were results presented in a clear, coherent fashion?
   ☒Yes
   ☐No
   ☐Unable to Code

Appendix E – Example Completed Coding Protocol for WoE A

Coding protocol used: Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C. & Innocenti, M. (2004). Quality indicators for group experimental and quasi-experimental research in special education. Exceptional Children, 71,149-164

Study: Brannon, D. & Dauksas, L. (2014). The Effectiveness of Dialogic Reading in Increasing English Language Learning in Preschool Children's Expressive Language. *International Research in Early Education,* 5(1), 1 – 10.

Essential and Desirable Quality Indicators for Group Experimental and Quasi-Experimental Research Articles and Reports

Essential Quality Indicators

Participants

1. Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?

   ☒Yes

   ☐No

   ☐Unable to Code

2. Was sufficient information provided to characterise the interventionists? Did it indicate whether they were comparable across conditions?

   ☒Yes

   ☐No

   ☐Unable to Code

Essential Quality Indicators

Implementation of the Intervention and Description of Comparison Conditions

3. Was the intervention clearly described and specified?

   ☒Yes

   ☐No

   ☐Unable to Code

4. Was the fidelity of implementation described and assessed?

☐Yes

☒No

☐Unable to Code

5. Was the nature of services provided in comparison conditions described?

☒Yes

☐No

☐Unable to Code

Essential Quality Indicators

Quality Indicators for Outcome Measures

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

☒Yes

☐No

☐Unable to Code

7. Were outcomes for capturing the intervention's effect measured at the appropriate times?

☒Yes

☐No

☐Unable to Code

Essential Quality Indicators

Data Analysis

8. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the limit of analyses in the study?

☒Yes

☐No

☐Unable to Code

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

☒Yes

☐No

☐Unable to Code

Desirable Quality Indicators

1. 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

   ☐Unable to Code

2. 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 (un)familiar to examinees across study conditions?

   ☐Yes

   ☒No

   ☐Unable to Code

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

   ☐Yes

   ☒No

   ☐Unable to Code

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

   ☒Yes

   ☐No

   ☐Unable to Code

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

   ☐Yes

   ☒No

   ☐Unable to Code

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

   ☒Yes

   ☐No

   ☐Unable to Code

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

   ☐Yes

   ☒No

   ☐Unable to Code

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

   ☒Yes

   ☐No

   ☐Unable to Code