

Case Study 1: An Evidence-Based Practice Report

Theme: Parent-

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).

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).

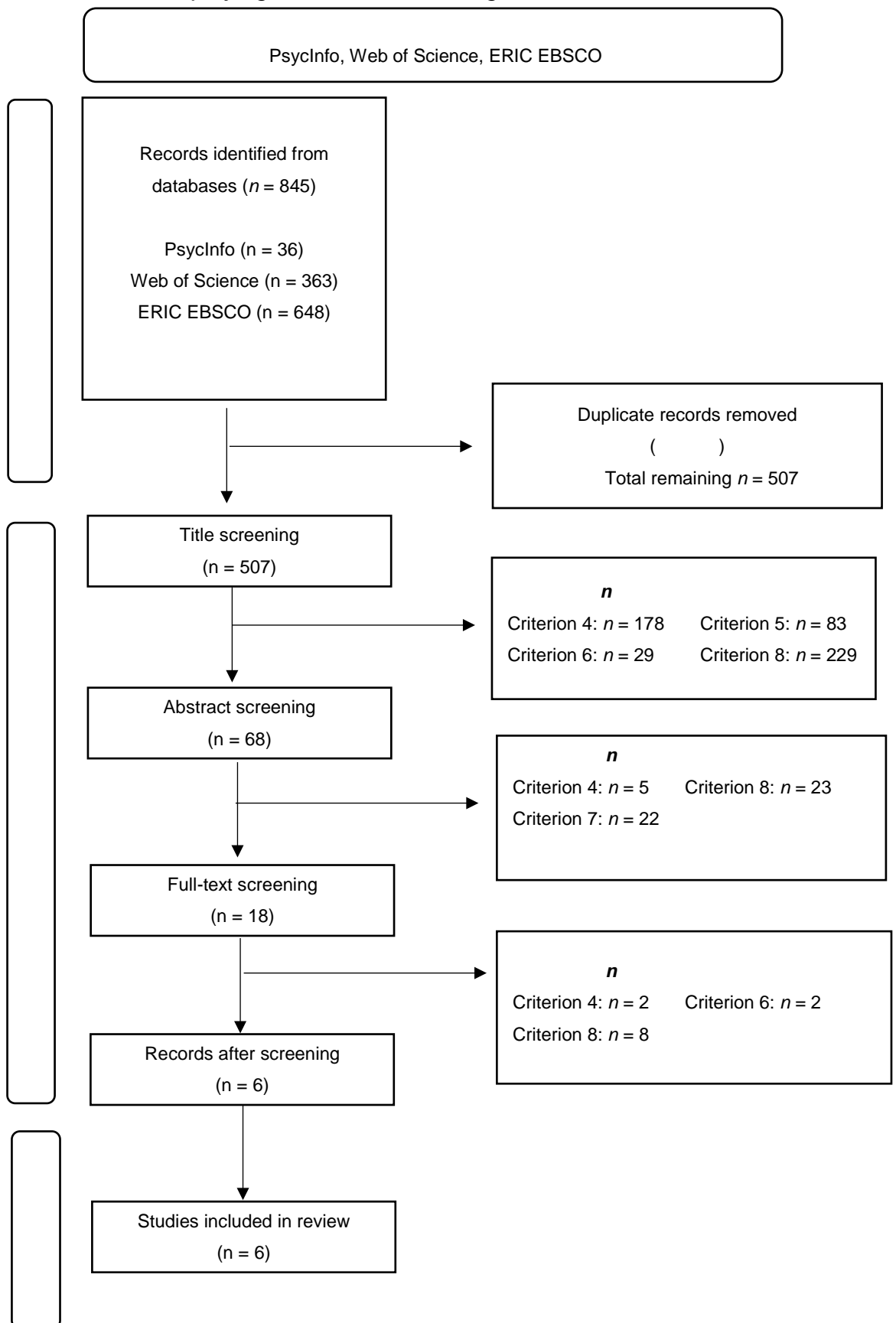
This review

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

Inclusion and Exclusion 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 near	

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 remaining 6 articles after screening. The final review

appraised 6 studies. Figure 1 displays an overview of this process. Table 2

oc. 1pd[.pl)6 (ay)14 (s)4 (a)]Jirv (nc)14(pl)6 (ay)14 (s)4 (i)16 (pnd[oc) a)]J[n]1016.3r0.0

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 ($n = 12$).

References for Final Studies Included in This Review

-
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.

(2018)	Three Head Start centres in New York.	126.	RCT.	Father parenting programme.	Children in the FSSP group showed improved auditory comprehension and expressive communication in comparison to the control group.
		Children: FSSP = 66% M, Waitlist = 69% M	Pre-test and post-test.	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.	
		FSSP mean age = 4.76; Waitlist mean age = 4.42 Fathers mean age: FSSP = 36.77; Waitlist = 35.25	FSSP (n = 64), Waitlist control (n = 62).		
(2020)	Bursa, Turkey.	80.			
		FSSP = 85% Hispanic; Waitlist = 89% Hispanic			
		Children: DR = 11 M, 9 F; Control = 12 M, 8 F			
		Parents: DR M, M t2020 = 4M, 16 F (30. i)-11.7 (z)-7 (e8s)-5 (p7.CID 52 >>BDC t)-6.6 (s)8.8 (:)-6.6 (Dey0 11.04 0 11.04 121.5e:5B03.56 T			

(2015)	Khayelitsha, Cape Town.	DR children mean age = 46.54 months.	91.	RCT	DR intervention group, control group.	Parents who had had the DR training reported significantly more words understood by their child.
		Children control: DR = 26 M, 16 F; Children DR = 33 M, 16 F		Pre-test and post- test.	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 vocalised by their child.
		Control Children mean age = 15.29 months; DR children mean age = 15.45 months.		DR group and control group.		
		Parent control mean age = 31.76 years Parent DR mean age = 33.35 years				

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.

Total Weight of Evidence Ratings for Studies Included in this Review

Brannon & Dauksas (2014)	1 (Low)	2.25 (Medium)	2.66 (High)	1.97 (Medium)
--------------------------	------------	------------------	----------------	------------------

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).

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.

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, 2014), 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

therefor received a lower WoE B rating for this category. All studies in this

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.

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

	<i>d</i>) *	
*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

d) *

*d) **

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.

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

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-

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

rs461660lx9R)6.3B(4a),10ar2(i)18:0004.1w2

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., Simo4od2Ni. S u5 (he v17 (eno(R.H)6 (ow)11 0.0

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)*,

Preschoolers with Autism Spectrum Disorder. *International Journal of*

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

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

1	Cohen, L., Kramer-Via, L. & Frye, N. (2012). Implementing Dialogic Reading with Culturally, Linguistically Diverse Preschool Children. <i>Research-to-Practice Journal for the Early Childhood Field</i> , 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. <i>Trials</i> , 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 Underprivileged Pre-School Children's Expressive Language Ability. <i>Script Journal: Journal of Linguistics and English Teaching</i> , 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. <i>Reading Research Quarterly</i> , 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. <i>Procedia-Social and Behavioural Sciences</i> , 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. <i>Teaching Exceptional Children</i> , 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

<p>spectrum disorder. <i>Journal of Early Intervention</i>, 40(1), 363 – 379.</p>	<p>researcher, not by parents as per the review question</p>
<p>8 Ganotice Jr, F. A., Downing, K., Mak, T., Chan, B., & Lee, W. Y. (2017). Enhancing parent-child relationship through dialogic reading. <i>Educational Studies</i>, 43(1), 51-66.</p>	<p>Criterion 6: This study investigated the impact of dialogic reading on parent-child relationships, not on child language development.</p>
<p>9 Rahn, N. L., Coogle, C. G., & Storie, S. (2016). Preschool children’s use of thematic vocabulary during dialogic reading and activity-based intervention. <i>The Journal of Special Education</i>, 50(2), 98-108.</p>	<p>Criterion 8: Dialogic Reading</p>

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.

Rating Criteria for WoE A

3 (High)	The study meets at least 9 of the essential criteria <i>and</i> at least 5 of the desirable criteria.
2 (Medium)	The study meets at least 9 of the essential criteria <i>and</i> at least 1 of the desirable criteria.
1 (Low)	The study meets at least 6 of the essential criteria.

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

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).

'Outcome measures

Criteria for WoE B Rating

Research

Overall WoE B Ratings for Included Studies

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

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.

WoE C Criteria and Rationale

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

to measure a skill relating to language.

1 – Only one area of language development has been assessed.

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.

Overall Ratings for WoE C for studies included in this Review

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)

Positive parenting (Dyadic Parent
Child Interaction Coding System)

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

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?

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

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

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 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 measures 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

