Journal of Research in Reading
UKLA The United Kingdom Literacy Association
Journal of Research in Reading, ISSN 0141-0423
Volume 00, Issue 00, 2022, pp 1–16
DOI:10.1111/1467-9817.12407

The story so far: A systematic review of the dialogic reading literature

Claire Pillinger
Centre for Teachers and Teaching Research, UCL Institute of Education, London, UK

Emma J. Vardy
Department of Psychology, Nottingham Trent University, Nottingham, UK

Background: Dialogic reading (DR; Whitehurst et al., 1988) is an evidence-based intervention that promotes children's active participation in shared reading (Towson, 2016; Urbani, 2020; WWC, 2007, 2010). Since the development of DR, there has been a proliferation of studies evaluating the conditions and populations with which it is effective. However, to date, there has not been a systematic review of the literature focusing specifically on the impact of DR on the literacy and non-literacy skills of children under 10 years old. As DR research evolves, it is important that a review of the existing literature is undertaken to capture these advances and guide future research.
Methods: Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol, the aim of this review is to systematically explore, synthesise and critically evaluate the extant literature. A systematic search of electronic databases identified 46 relevant studies, and the overall methodological quality of the studies was assessed using the MMAT.
Results: Findings are organised according to sample and population, country of origin and setting, programme duration, language and literacy outcomes, social-emotional and other cognitive outcomes, impact and effect sizes to provide overview and insight into where and with whom DR is most effective.
Conclusions: The review findings suggest DR can positively impact a wide range of language and literacy skills for children under 5 years. There is some evidence that DR can have positive effects on enjoyment of reading, reading motivation, parental–child attachment, parental confidence and stress. However, the extant research is subject to limitations, and more methodologically robust research is needed to enable thorough assessment of the conditions in which DR is most effective.

Keywords: Dialogic reading, Early literacy, Shared reading, Family literacy, Literacy practices

Highlights

What is already known about this topic

- Shared reading can have a positive impact on a wide range of language skills for young children.

- DR is an easy-to-administer interactive shared book reading intervention for parents and educators.
- DR can have positive effects on the language and literacy development of young children.

What this paper adds

- Previous literature reviews have been somewhat limited in scope, utilising a limited number of search engines, comprising a small number of studies and focusing solely on language outcomes. This is the first comprehensive systematic review, focusing on the impact of DR on language and literacy outcomes, social-emotional and other cognitive outcomes for children under 10 years.
- Provides a summary of the extant research on DR (based on the Whitehurst et al., 1988 model).
- Provides a quality appraisal of the extant DR literature.

Implications for theory, policy or practice

- DR can have positive effects on the language, literacy and social-emotional development of young children (<5 years)
- More methodologically robust research is needed to identify the practical and/or theoretical importance of the DR intervention (e.g. calculation of reporting of effect size) and the effects of DR on complex language skills.
- DR could be considered a useful intervention for increasing parental engagement in shared book reading. It may be the case that DR's more structured approach places fewer demands on parents who are less confident reading with their child and therefore provides a useful starting point for encouraging parental engagement in joint storybook reading interactions. The increased exposure to books in the home, in turn, facilitates language development for the child.

'Shared book reading' encompasses various read-aloud methods and book-related activities that support children's language and literacy development. Children acquire important literacy skills through shared book reading including alphabetic knowledge, phonological awareness, print awareness, knowledge of syntax, semantics, pragmatics and narrative structure (Senechal & LeFevre, 2001).

The influence of shared book reading on children's early language and literacy learning is supported by extensive research (Bus et al., 2007; Saracho & Spodek, 2010). Pre-test–-post-test design studies show that children engaging in shared book reading in the home score higher on post-intervention reading assessments than children who did not experience storybook reading (Kotaman, 2008; Roberts, 2008; Smetana, 2005). Although the benefits of shared reading are well established, adults do not always know how best to support children (Fielding-Barnsley & Purdie, 2003). A number of shared reading interventions have therefore been developed to help adults maximise the benefits of the shared reading experience.

One such intervention is dialogic reading (DR). DR is an evidence-based intervention developed by Whitehurst et al. (1988) that promotes children's active participation in shared reading interactions. In contrast to traditional styles of shared book reading where an adult reads the book and the child passively listens, DR encourages the child to become the storyteller and the adult an active listener who provides assistance as required, facilitates, expands and responds to the child's verbalisations.

DR is based on three principles: (i) *evocative techniques* that encourage the child to actively participate in reading and practice language, (ii) the use of *informative feedback* for the child regarding correct language use and (iii) *progressive change* where the adult adapts their reading style to the child's developing linguistic abilities. In DR training, adults are taught a set of standardised procedures that encourage the child to engage in dialogue about the book they are reading. To facilitate this, adults are taught a set of prompts represented by the acronyms PEER and CROWD. The PEER sequence helps adults remember the specific order of DR techniques and includes PROMPT, EVALUATE, EXPAND and REPEAT, which specifically target vocabulary and comprehension skills. However, the CROWD prompts, which includes COMPLETION, RECALL, OPEN-ENDED questions, WH questions and DISTANCING, encourage the child to express themselves and develop existing language abilities through practising linguistically enhanced dialogue (Senechal, 2006).

Since the inception of DR in the late 1980s, there has been a proliferation of studies evaluating the conditions and populations with which DR is effective. Although reviews of the literature have been previously undertaken (Towsen et al., 2017; What Works Clearing House [WWC], 2007, 2010), they are somewhat limited in scope. For example, Towsen and Gallagher (2014) evaluated the evidence base of DR across early childhood settings, with specific focus on fidelity. Thirty peer-reviewed articles were identified using two search engines (PsycINFO, ERIC) and coded for participant characteristics, setting, outcomes, study rigour, fidelity of training and implementation. The review revealed that a wide range of child outcome measures were used, including standardised language assessments (e.g. PPVT and EOWPVT), informal language assessments [e.g. mean length of utterance (MLU) and verbal participation], standardised literacy assessments [e.g. concepts about print (CAP)] and informal emergent literacy assessments (e.g. rhyme awareness). Moreover, 27 studies employed researcher developed tools designed to directly assess targeted outcomes of intervention, only eight studies tracked changes in adult behaviour related to DR, and similarly, there was high variability in the type of measures used.

Towsen et al. (2017) also report that DR interventions were primarily implemented within home and school settings by caregivers and educators, 70% of studies involved typically developing children or those at risk for later deficits and few studies (n = 8) involved children with an identified disability. It was also found that although training procedures were clearly explained in all studies, none reported the fidelity of those practices. Towsen et al. (2017) therefore concluded that it is difficult to determine for whom DR is effective because gains in oral language cannot be directly attributed to DR without adequate evidence that the intervention was implemented as designed.

The WWC (2007) review consisted of five studies that examined intervention effects on children's oral language and phonological processing. The children in the sample were aged 2–5.5 years, mostly from economically disadvantaged families. Analysis of the five studies revealed that DR had statistically significant effects on oral language skills

including receptive, expressive vocabulary, novel vocabulary and verbal fluency; however, there were no discernible effects on phonological processing.

WWC (2010) included two studies that examined the effects of DR on the communication and language competencies of children aged 3–6 years with disabilities. One study by Dale et al. (1996) did not find statistically significant differences between DR and control groups on receptive or expressive vocabulary, MLU, number of utterances or lexical diversity. Dale et al. (1996) found a statistically significant difference favouring DR group children on lexical diversity, but no statistically significant differences between groups on MLU and number of child utterances. Consistent with the findings of Towsen et al. (2017), the WWC concluded that DR had potentially positive effects on communication and language competencies for children with disabilities, but more research involving children with disabilities was needed to fully understand the impact.

Although the previous reviews are insightful, they utilised a limited number of search engines, comprise a small number of studies and focus solely on language outcomes. A comprehensive and systematic review of the literature focusing on the impact of DR on language, literacy, social-emotional and other cognitive outcomes for children under 10 years is therefore needed. As DR research continues to evolve, it is important that a review of the DR literature is undertaken to capture these advances and guide future research. Thus, the aim of this literature review is to systematically explore, synthesise and critically evaluate the extant academic literature on DR. The results from the studies will be summarised, providing researchers and educators with up-to-date knowledge of the area, identifying conditions in which DR is most effective and by both critically evaluating and highlighting gaps in the extant literature, guide future research.

Method

Search strategy: Sources of literature

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Moher et al., 2009) protocol was used to guide the conduct and reporting of this review. A systematic search of relevant electronic bibliographic databases was undertaken including Academic Search Complete, PsycArticles, PsycINFO, ERIC, Web of Science and PubMed. The first search was completed by 11 September 2019, and a second search was carried out on 3 July 2021. The reference list of included studies was also searched and considered in line with the inclusion/exclusion criteria.

Search strategy: Search terms and inclusion criteria

Search terms were generated through discussion with review authors and consideration of terms used in previous literature reviews. In order to capture all the available research, the following broad search terms were applied to all databases: *dialogic reading, interactive shared reading*.

For the academic literature, pre-defined inclusion and exclusion criteria were applied to studies retrieved through the searches. The following inclusion criteria were applied: Texts had to (i) include participants aged birth—10 years (this age range was selected as research suggests that parent reading significantly declines after the age of 10 years)s (Merga, 2017; Scholastic Kids & Family Reading Report, 2018), (ii) be published in English language,

(iii) in peer-reviewed journal articles, (iv) be original empirical studies, (v) include an original or adapted form of the DR programme developed by Whitehurst et al. (1988) and (vi) be published between 1988 and 2021. Studies obtainable within 2 weeks of being identified were included in the review. For this part of the search, articles were excluded if they were books, newspapers articles, conference papers, thesis and unpublished articles. The inclusion criteria are designed to capture all research designs to ensure the review includes the largest sample of DR research published to date. Although RCTs are considered the gold standard for research, focusing solely on this design for a systematic review risks excluding a large proportion of studies and would present an unrepresentative summary and evaluation of the DR literature.

Search strategy: Screening

Each database was searched separately using the search terms above. A double screening approach was adopted to ensure reliability, validity and elimination of bias in the selection of studies. The first author screened the titles and abstracts of all identified studies to determine whether they met inclusion criteria, the second author independently screened 50% of the identified studies, and 100% agreement was reached. Duplicates from previous searches were removed. Articles were also excluded if it was clear that they did not meet the above criteria. However, articles were retained if they met or if it was unclear whether they met the inclusion criteria, and the full papers were then read. A hand search of the reference section of all identified articles was also undertaken in order to identify all possible relevant articles. See Figure 1 for the search strategy for the academic literature.

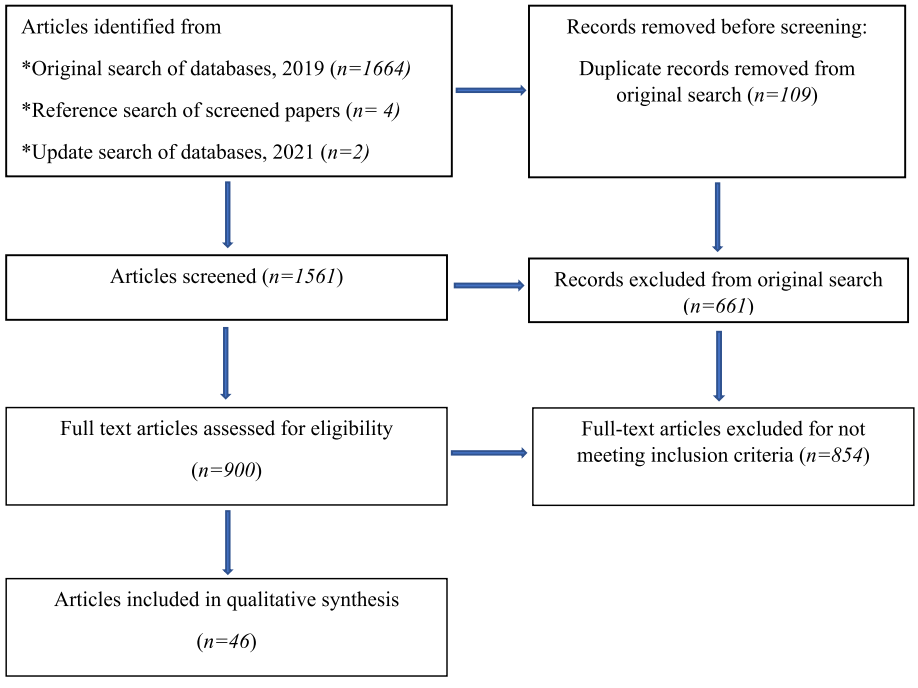

Figure 1. Search strategy based on the PRISMA model. [Colour figure can be viewed at wileyonlinelibrary.com]

Study selection

A total of 1664 articles from the systematic search of the databases were identified following the first search. Duplicate articles (n = 109) were removed, leaving 1555 records for screening. Following initial screening of the title and abstract, 661 articles were removed, leaving 894 full-text articles for evaluation. Subsequent full-text screening resulted in 854 articles being removed as they did not meet the inclusion criteria. An additional four articles were identified from hand searches, and a second search in July 2021 identified two articles for inclusion. Forty-six articles comprising 46 independent studies were therefore retained for quality appraisal and inclusion in the review. A summary of the findings from the papers can be found here: https://discovery.ucl.ac.uk/id/eprint/10150515/.

Data extraction

A double extraction approach was adopted. The studies identified were divided between the authors. Each author independently extracted data from the allocated studies using an electronic spreadsheet. Both authors then cross-referenced and checked 50% of the data extraction tables, and 100% agreement was reached. Extracted data included sample characteristics, recruitment strategy, intervention characteristics (duration/frequency/setting), aim and methodology, outcome measures (literacy tests/observation schedules/fidelity measures used) and outcomes/findings relevant to the review.The studies included in this review utilised different outcomes measures, time frames and analytic strategies; therefore, the data extracted did not allow for meta-analysis due to heterogeneity of outcomes. The data for each study have been drawn together from the data extraction table to present a narrative synthesis of results.

Quality appraisal: Mixed Methods Appraisal Tool

This systematic review aimed to include a wide range of study design to gain a richer understanding of the impact of contextual factors and explore the diversity of effect across studies. In the quality appraisal section of this review, findings from the RCTs conducted to date will be considered. In addition, the Mixed Methods Appraisal Tool (MMAT; Pluye et al., 2011) was selected to allow a more fine-grained quality analysis of the studies identified. The MMAT is a comprehensive evaluation tool that allows for the concomitant appraisal of the methodological quality of qualitative, quantitative and mixed methods designs. Both authors assessed the quality of all studies meeting inclusion criteria and tabulated scores for each review before awarding an overall quality star rating. Both authors are experienced in conducting all types of study and have good understanding of the designs used so were able to accurately and fairly decide if the MMAT criteria were met. Studies were rated using a star system, with four stars indicating the highest methodological quality. Within the narrative discussion, studies rated as zero or one star are considered low-quality studies, whereas those rated as three or four stars are regarded as high-quality studies. The first author completed quality appraisal checks on 50% of the paper, and 50% were completed by the second author. The first author then checked 10% of the papers appraised by the second author and vice versa. There were no areas of disagreement in rating of quality assessment.

Results

As can be seen, the efficacy and impact of the DR programme has been evaluated in a range of settings and countries, with various populations using a variety of outcome measures. The research is summarised later and has been separated into sample and population, country of origin and setting, programme duration, language and literacy outcomes, social-emotional and other cognitive outcomes, impact and effect sizes. This will provide insight into where and with whom DR is most effective; however, it is important that the methodological rigour of the studies should be considered when evaluating the impact of the reported findings. Quality appraisal and limitations of the extant research are also discussed along with suggestions for future research.

Sample and population

Forty-six papers were included in this review, representing a total sample of 2693 participants aged 2–9 years old (the modal age range = 2–5 years). The 36 articles that included data on gender revealed that there were 1253 male child participants and 1035 females. In contrast, with the exception of one study (Pillinger & Wood, 2013), 80–100% of the adult participants were female. Sample sizes ranged from 3 to 324, with most studies including a sample of 30–40 participants, which is considered small for quantitative research given the homogeneity of the sample, primary outcome measures, the calculated margin of error considered appropriate for the question being studied and the anticipated attrition. (Burmeister & Aitken, 2012; Lenth, 2001).

The studies included typically developing children, dual-language learners, deaf and hard-of-hearing children, those with or at risk of language delays, children with disabilities, special educational needs and developmental disorders such as autism.

Twelve studies included data on ethnicity and revealed a wide range of ethnicities were represented across the 46 articles including White (286), Black (46), Hispanic (37), non-Hispanic (8), non-Hispanic Black (13), Asian Pacific (1), Asian (16), African American (549) and Latin American (46).

Only 22 studies provided sufficient information on the socio-economic background of their sample, indicating that most participants were of low socio-economic status (*n = 18*).

Country of origin and setting

The impact of DR on young children's language abilities has been demonstrated in many countries. In this review, most studies were based in the United States (27), followed by Hong Kong (5), Canada (3), Australia (2), Turkey (2), the United Kingdom (2), Egypt (1), Bangladesh (1), Brazil (1), Mexico (1) and South Africa (1). DR was found to be effective when delivered in English and when adapted to other languages, for example, Arabic (Elmonayer, 2013), Bangla (Opel et al., 2009), Chinese (Chow et al., 2008, 2010; Chow & McBride-Chang, 2003) and Spanish (Cohen et al., 2012; Huennekens & Xu, 2016; Jimenez et al., 2006; Towsen & Gallagher, 2014).

The papers included in this review revealed that the DR intervention has been implemented in a range of settings including day care, nurseries, community centres, libraries, schools and within the home.

Programme duration

Only 28 studies included information on the frequency and/or duration of the sessions. These studies revealed that DR interventions ranged from 5 to 20 min per day, a minimum of twice a week to 7 days a week. The modal dosage for DR was 5–15 min per day, for 6 weeks. To date, no study has investigated optimal dosage for DR interventions.

The duration of DR programmes ranged from 2 weeks to 1 school year, with most interventions lasting 6–8 weeks. The conclusion from the studies examined in this review is that short-term DR interventions can have immediate effects on both parental reading behaviour, children's language abilities and social development.

Data on the longer term impact of DR on literacy outcomes are somewhat tentative as few studies (n = 7) included follow-up testing (Fielding-Barnsley & Purdie, 2003; Huebner, 2000a; Huebner & Payne, 2010; Pillinger & Wood, 2014; Sim et al., 2014; Whitehurst et al., 1988, 1994). In the seven studies, the follow-up interval ranged from 6 weeks to 2 years, with most including follow-up 3–9 months after the intervention was terminated.

Findings revealed that DR had long-term effects on parental reading behaviours as those trained in DR techniques continued to use the techniques up to 2 years later (Huebner, 2000a; Huebner & Payne, 2010). Three studies reported that positive effects on expressive language were maintained 3–9 months later (Huebner, 2000a; Whitehurst et al., 1988, 1994), and there was also some evidence that gains in receptive language and MLU were maintained (Whitehurst et al., 1988), along with writing vocabulary, word reading (Pillinger & Wood, 2014), final consonant recognition (Fielding-Barnsley & Purdie, 2003), picture naming and rhyme (Sim et al., 2014). The longer term effects of DR on concepts about print (CAP) were mixed as two studies (Fielding-Barnsley & Purdie, 2003; Sim et al., 2014) reported that gains were maintained at 3 and 9 months follow-up, whereas Pillinger and Wood (2014) found that DR group gains on CAP during the intervention were not evident 3 months later.

Language and literacy outcomes

DR has been found to positively impact a wide range of important language and literacy skills, including receptive vocabulary (Fung et al., 2005; Kotaman, 2008; Pillinger & Wood, 2014; Simsek & Erdogan, 2015; Towsen & Gallagher, 2014; Valdez-Menchaca & Whitehurst, 1992), expressive vocabulary (Arnold et al., 1994; Hargrave & Senechal, 2000; Huebner, 2000a; Lever & Senechal, 2011; Opel et al., 2009; Rahn et al., 2016; Reese et al., 2010; Simsek & Erdogan, 2015; Valdez-Menchaca & Whitehurst, 1992; Whitehurst et al., 1988, 1994), phonological awareness (Chow et al., 2010; Elmonayer, 2013), rhyme recognition (Fielding-Barnsley & Purdie, 2003; Sim et al., 2014), word reading (Chow et al., 2010; Pillinger & Wood, 2014), CAP (Fielding-Barnsley & Purdie, 2003; Lonigan et al., 2013; Pillinger & Wood, 2013; Sim et al., 2014), character recognition (Chow et al., 2008, 2010), contextual knowledge (Lever & Senechal, 2011), MLU (Huebner, 2000a; Huebner, 2000b; Huebner & Meltzoff, 2005; Jimenez et al., 2006; Lonigan & Whitehurst 1998; Valdez-Menchaca & Whitehurst, 1992; Whitehurst et al., 1988) and language complexity (Huebner, 2000a; Valdez-Menchaca & Whitehurst, 1992).

However, it should be noted that the majority of DR studies conducted to date utilised a limited range of standardised oral language measures. The most common were the Peabody Picture Vocabulary Test Revised (PPVT-R; Dunn & Dunn, 1981), Expressive One Word Picture Vocabulary Test (EOWPVT; Gardner, 1981) and the Illinois Test of Psycholinguistic Abilities (ITPA; Kirk et al., 1968), which are considered measures of 'simple language' (NELP, 2008). There are few DR studies that employ 'complex' or 'composite' language measures that assess skills like grammar and listening comprehension, which are considered 'stronger predictors of later reading achievement than measures of simple vocabulary' (NELP, 2008, p. 157). The only study in this review to assess listening comprehension was that of Lonigan et al. (1999) who allocated children aged 2–5 years from low-income backgrounds to an either DR, typical reading (TR) or a control group and found TR children outperformed DR group children. Other studies have looked at story comprehension, for example, Pfeiffer-Flores et al. (2014) investigated the effects of DR on textual comprehension of a children's novel with three Brazilian children and found that comprehension scores increased more for DR group children. Similarly, Vally et al. (2015) found that DR children identified a greater number of items on a measure of language comprehension based on PPVT-R, compared with control group; however, after controlling for the effect of baseline performance, the difference was not significant. Findings regarding the contribution of DR to higher level language and literacy abilities are therefore limited and uncertain. More research is needed to ascertain the true extent of the impact of DR on different types and levels of comprehension.

Social-emotional and other cognitive outcomes

In addition to literacy ability, DR has also been found to enhance engagement (Fleury, 2013), enjoyment (Huebner, 2000b; Pillinger & Wood, 2014), interest and frequency of reading (Beschorner & Hutchison, 2016; Chow et al., 2008; Chow & McBride-Chang, 2003; Huebner, 2000b; Huebner & Payne, 2010; La Cour et al., 2013), improve communication between parent and child (Ganotice et al., 2017), enhance sustained attention (Vally et al., 2015), strengthen the bond between parent and child, increase satisfaction with parenting (Ganotice et al., 2017; Huebner, 2000b), lower parental stress (Huebner, 2000a) and increase confidence of caregiver (La Cour et al., 2013). A range of standardised and non-standardised measures have been used to assess these outcomes including video/in-person observation with behavioural coding, interviews with parents, questionnaires, the Parent–Child Relationship Inventory, Parent Stress Index and the Preschool Reading Attitudes Scale. The heterogeneity of measures renders comparison between studies somewhat problematic. Also, a lot of the measures have been used to assess parental satisfaction, changes in reading behaviours are self-report, and without additional evidence or observation, it is difficult to ascertain whether parents' response were an honest and accurate reflection of their home reading practices or if they were in some way biased to meet the perceived expectations of the research team/intervention staff.

Impact and effect sizes

Although significant differences between intervention and control/alternative treatment groups are reported for DR interventions, the statistical significance simply indicates

whether obtained results were likely to have occurred by chance. It is the effect size for an intervention that indicates the relative magnitude of difference between group means and whether an intervention has educational value. Closer inspection of the current DR research reveals that very few of the published studies calculated and reported the effect size (n = 8) (Blom-Hoffman et al., 2007; Chow & McBride-Chang, 2003; Chow et al., 2008, 2010; Huennekens & Xu, 2016; Lonigan & Whitehurst, 1998; Pillinger & Wood, 2013; Whitehurst et al., 1994). It is only in subsequent meta-analytic studies that mean oral language effect sizes for DR interventions have been made explicit. Moreover, the meta-analytic reviews reveal that effect sizes for DR interventions based on the Whitehurst et al. (1988) method are highly variable (Schickedanz & McGee, 2010). For example, an NELP (2008) meta-analysis of 15 studies revealed moderate effects on oral language, whereas a meta-analysis by Schichedanz & McGee (2010: 326) concluded 'dialogic reading has a differential effect between younger and older pre-schoolers … the effects of shared reading interventions seemed more pronounced for younger than older pre-schoolers'.

Given the current findings, it could be concluded that the educational value of the Whitehurst et al. (1988) programme is modest and more studies need to report effect sizes so that the true educational value of the DR intervention can be ascertained.

Quality of the extant DR literature

Randomised control trials (RCTs) are considered the gold standard of educational research. However, only two of the 46 studies included in this review were RCTs (Sim et al., 2014; Vally et al., 2015). Sim et al. (2014) randomly assigned families to one of three groups: DR, dialogic reading with the addition of print referencing (DR + PR) or an attention-matched control group for 8 weeks. Results revealed that although there were no significant differences between groups on measures of receptive vocabulary, word completion, alphabet knowledge; children in the DR and DR + PR group showed greater improvement on measures of rhyme, rapid naming ability and CAP than the control group. There were no significant differences between the DR and DR + PR groups on any of the outcome measures. At follow-up, children in the DR and DR + PR groups maintained the effects of the intervention compared with the children in the control group for CAP only.

Vally et al. (2015) investigated the impact of DR on child language and attention amongst 91 infant–mother dyads in an impoverished South African community. After 8 weeks of individual DR sessions at home and groups sessions at school, DR group parents reported that their infants could both understand and vocalise a significantly greater number of words (as measured by the MacArthur-Bates CDI). This difference, after controlling for baseline scores, was highly significant. DR group children's comprehension scores improved more than the control group; however, after controlling for baseline scores, the difference was not significant. DR children showed a substantial increase in sustained attention (ECVT), whereas those in the control group showed no change; this difference was significant after controlling for baseline scores.

Although both RCTs are limited by the absence of intervention fidelity monitoring, the results do suggest that DR can positively impact rapid naming ability, rhyme, CAP, sustained attention and expressive and receptive vocabulary.

MMAT appraisal

Of the 46 studies, two met the MMAT criteria for qualitative research, 23 as randomised-quantitative, eight as quantitative non-randomised, 11 quantitative-descriptive and two mixed methods. Seven studies were considered high-quality 4* papers, nine 3*, 21 as 2*, five as 1* and five as 0* (see column 1 of the table: https://discovery.ucl.ac.uk/id/eprint/10150515/). This suggests that the majority of the DR studies to date are considered poor quality using the MMAT rating tool. The main methodological concerns in the extant data are the small sample sizes and the lack of follow-up after the termination of the study, which raises questions about whether there is sufficient time for the outcome to occur and long-term impact to be assessed. Of the three qualitative studies, only one considered how findings relate to the context (e.g. the setting in which the intervention was delivered or data collected), and none considered whether and to what extent researcher influence impacts upon outcomes. In the 25 quantitative-randomised studies, although reference was made to participants being randomly allocated to treatment groups, just 10 provided adequate information about the process of randomisation, and only four explained the process of allocation concealment. In the non-randomised quantitative studies, the main weaknesses were the lack of information regarding recruitment processes. Just two of the nine studies provided information about measures taken to minimise selection bias, and less than half ($n = 4$) reported that efforts had been made to minimise/control group differences (e.g. whether participants were recruited in a way to minimise selection bias or group differences were controlled for).

In the majority of DR studies conducted to date, few studies provide adequate demographic data regarding the age, socio-economic background and ethnicity of the sample. This inevitably raises concerns about the generalisability of findings and the external validity of DR research as a whole as little is known about the efficacy of DR amongst different strata of the population.

In half ($n = 4$) of the quantitative-descriptive studies, it was not clear whether the sample was representative of the population being studied, and in eight of the 11 studies, it was not clear whether the measures used were of suitable validity as they employed bespoke measures developed by the researchers to measure targeted outcomes or were adapted to the native language of the participants.

For the mixed methods studies, evaluation according to the MMAT criteria revealed that in most papers, the qualitative and quantitative data were poorly integrated, and none of the included studies gave appropriate consideration to the limitations associated with triangulation design and the integration.

Discussion

The extant research suggests DR can have a positive impact on a wide range of language skills for young children. Across the studies included in this review, gains were most consistently recorded on measures of receptive and expressive language. There is also some evidence that DR can have positive effects on social-emotional outcomes, including enjoyment of reading, reading motivation, parent–child attachment, parental confidence and stress. Moreover, follow-up studies suggest DR intervention can have long-term impact, with gains in children's receptive and expressive language maintained up to 9 months later and changes in parental behaviour evident after 2 years.

Caution is however needed when interpreting the findings of the extant literature for a number of reasons. Firstly, this review highlighted that just 27 studies included utilised measures of intervention fidelity. Of the studies that did include intervention fidelity measures, there was little consistency in what methods were used. Methods used included reading logs, videotape or observations of a few sessions or audio recording of sessions. The fidelity measures were primarily self-report; thus, the extent to which they are an accurate reflection of true implementation is questionable. Seventeen of the 27 studies reported good fidelity; however, eight of the studies reported that the data were incomplete, so precluded analysis of whether or not the intervention was implemented as the researcher intended.

Lack of control groups was also a major limitation in the literature reviewed. Just 23 of the 44 studies included in the review had a control group. Thus, it cannot be ascertained whether the changes that occurred were the direct result of the DR intervention, extraneous variables, the Hawthorne effect or simple passage of time and would have occurred anyway through participation in the regular school curriculum.

The current review is also subject to limitations. To the authors' knowledge, the MMAT rating tool has not previously been used to evaluate educational research. Therefore, although it is a valuable tool in enabling consistent evaluation of different methodologies and popular in other areas of psychological research, its sensitivity and appropriateness to evaluating methodological rigour of educational research warrant further investigation.

Despite these limitations, DR could be considered a useful intervention for increasing parental engagement in shared book reading. It may be the case that DR's more structured approach places fewer demands on parents who are less confident reading with their child and therefore provides a useful starting point for encouraging parental engagement in joint storybook reading interactions. The increased exposure to books in the home in turn facilitates language development for the child.

Future research

This review has identified a number of weaknesses in the extant literature that future research should aim to address. Firstly, only three studies in this review assessed the impact of DR on more complex measures of language such as grammar and comprehension, and the findings of these were mixed. Thus, future research needs to include more complex language measures as to gain a stronger indication of the impact of DR intervention for enhancing language development and later language skills.

Relatedly, rather than a reliance on self-report or observation measures, more robust, valid and reliable measures or a triangulation approach are needed to capture the true extent of the impact of DR on socio-emotional outcomes such as reading enjoyment and parental confidence. More valid and reliable measures of intervention fidelity would also help researchers identify the optimal conditions for DR and barriers to implementation.

Future studies should also take care to include control groups, ensure group differences are controlled for and make sure processes of randomisation are fair and clear to enable replication so that differences between groups post-intervention can be confidently attributed to intervention effects and not pre-existing differences between groups.

Future DR research should also seek to calculate effect sizes so that the practical and/or theoretical importance of the DR intervention can be ascertained. The inclusion of a follow-up testing phase after termination of the intervention would also provide a clearer

indication of the long-term impact of the intervention and whether meaningful changes in parent–child behaviour are sustained once the intervention ceases. Only Huebner and Payne (2010) looked at the long-term impact of DR instruction, and although they report that after a 2-year follow-up, DR instruction had long-lasting impact on reading behaviour, more studies are needed to corroborate this finding.

Finally, larger and more varied sample including male caregivers, teachers, teaching assistants and children (e.g. peer-assisted) will enable a more comprehensive evaluation of the efficacy of DR and the populations with which it is most effective.

Addressing the aforementioned limitations should help create a more robust data set for the efficacy of DR and enable educators to make an informed decision about when, where and with whom this intervention can be implemented to maximise impact.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology*, 86(2), 235–243. https://doi.org/10.1037/0022-0663.86.2.235

Beschorner, B., & Hutchison, A. (2016). Parent education for dialogic reading: Online and face-to-face delivery methods. *Journal of Research in Childhood Education*, 30(3), 374–388. https://doi.org/10.1080/02568543.2016.1178197

Blom-Hoffman, T., O'Neil-Pirozzi, R., Cutting, J., & Bissinger, E. (2007). Instructing parents to use dialogic reading strategies with preschool children. *Journal of Applied School Psychology*, 23(1), 117–131. https://doi.org/10.1300/J370v23n01_06

Burmeister, E., & Aitken, L. M. (2012). Sample size: How many is enough? *Australian Critical Care: Official Journal of the Confederation of Australian Critical Care Nurses*, 25(4), 271–274. https://doi.org/10.1016/j.aucc.2012.07.002

Bus, A., de Jong, M., & Van Ijzendoorn, M. (2007). Social aspects in language and literacy learning: progress, problems, and interventions. In O. Saracho & B. Spodek (Eds.), *Contemporary perspectives on social learning in early childhood education* (pp. 243–258). Information Age.

Chow, B. W.-Y., & McBride-Chang, C. (2003). Promoting language and literacy development through parent–child reading in Hong Kong pre-schoolers. *Early Education and Development*, 14(2), 233–248. https://doi.org/10.1207/s15566935eed1402_6

Chow, B. W.-Y., McBride-Chang, C., & Cheung, H. (2010). Parent-child reading in English as a second language: Effects on language and literacy of Chinese kindergarteners. *Journal of Research in Reading*, 33(3), 284–301. https://doi.org/10.1111/j.1467-9817.2009.01414.x

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

Cohen, L., Kramer-Vida, L., & Frye, N. (2012). Using dialogic reading as professional development to improve students' English and Spanish vocabulary. *NHSA Dialog: A Research-to-Practice Journal for the Early Childhood Field*, 15(1), 59–80. https://doi.org/10.1080/15240754.2011.636490

Dale, P., Crain-Thoreson, C., Notari-Syverson, A., & Cole, K. (1996). Parent-child book reading as an intervention technique for young children with language delays. *Topics in Early Childhood Special Education*, 16(2), 213–235. https://doi.org/10.1177/027112149601600206

Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test-revised*. American Guidance Service Inc.

Elmonayer, R. (2013). Promoting phonological awareness skills of Egyptian kindergarteners through dialogic reading. *Early Child Development and Care*, 183(9), 1229–1241. https://doi.org/10.1080/03004430.2012.703183

Fielding-Barnsley, R., & Purdie, N. (2003). Early intervention in the home for children at risk of reading failure. *Support for Learning*, 18(2), 77–82. https://doi.org/10.1111/1467-9604.00284

Fleury, V. P., Miramontez, S. H., Hudson, R. F., & Schwartz, I. S. (2013). Promoting active participation in book reading for preschoolers with Autism Spectrum Disorder: A preliminary study. *Child Language Teaching and Therapy*, 30(3), 273–288. https://doi.org/10.1177/0265659013514069

Fung, P. C., Chow, B. W., & McBride-Chang, C. (2005). The impact of a dialogic reading program on deaf and hard-of-hearing kindergarten and early primary school-aged students in Hong Kong. *Journal of Deaf Studies and Deaf Education*, 10(1), 82–95. https://doi.org/10.1093/deafed/eni005

Ganotice, F., Downing, K., Mak, T., Chan, B., & YipLee, W. (2017). Enhancing parent-child relationship through dialogic reading. *Educational Studies*, 43(1), 51–66. https://doi.org/10.1080/03055698.2016.1238340

Gardner, M. (1981). *Expressive one word picture vocabulary test*. Academic Therapy.

Hargrave, A., & Senechal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, 15(1), 75–90. https://doi.org/10.1016/S0885-2006(99)00038-1

Huebner, C. (2000a). Promoting toddlers' language development through community-based intervention. *Journal of Applied Developmental Psychology*, 21(5), 513–535. https://doi.org/10.1016/S0193-3973(00)00052-6

Huebner, C. (2000b). Community-based support for preschool readiness among children in poverty. *Journal of Education for Students Placed at Risk*, 5(3), 291–314. https://doi.org/10.1207/S15327671ESPR0503_6

Huebner, C., & Meltzoff, A. (2005). Intervention to change parent-child reading style: A comparison of instructional methods. *Applied Developmental Psychology*, 26(3), 296–313. https://doi.org/10.1016/j.appdev.2005.02.006

Huebner, C., & Payne, K. (2010). Home support for emergent literacy: Follow-up of a community-based implementation of dialogic reading. *Journal of Applied Developmental Psychology*, 31(3), 195–201. https://doi.org/10.1016/j.appdev.2010.02.002

Huennekens, M., & Xu, Y. (2016). Using dialogic reading to enhance emergent literacy skills of young dual language learners. *Early Child Development and Care*, 186(2), 324–340. https://doi.org/10.1080/03004430.2015.1031125

Jimenez, T., Filippini, A., & Gerber, M. (2006). Shared reading within Latino families: An analysis of reading interactions and language use. *Bilingual Research Journal*, 30(2), 431–452. https://doi.org/10.1080/15235882.2006.10162884

Kirk, S., McCarthy, J., & Kirk, W. (1968). *Illinois test of psycholinguistic abilities*. University of Illinois Press.

Kotaman, H. (2008). Impacts of dialogical storybook reading on young children's reading attitudes and vocabulary development. *Reading Improvement*, 43, 134–145. https://www.researchgate.net/publication/284163177

La Cour, M., McDonald, C., Tissington, L., & Tomason, G. (2013). Improving pre-kindergarten children's attitude and interest in reading through a parent workshop on the use of dialogic reading techniques. *Reading Improvement*, 50, 1–11. https://projectinnovation.com/reading-improvement

Lenth, R. V. (2001). Some practical guidelines for effective sample size determination. *The American Statistician*, 55(3), 187–193. https://doi.org/10.1198/000313001317098149

Lever, R., & Senechal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology*, 10(1), 1–24. https://doi.org/10.1016/j.jecp.2010.07.002

Lonigan, C., Anthony, J., Bloomfield, B., Dyer, S., & Samwell, C. (1999). Effects of two shared reading interventions on emergent literacy skills of at-risk pre-schoolers. *Journal of Early Intervention*, 22(4), 306–322. https://doi.org/10.1177/105381519902200406

Lonigan, C., Purpura, D., Wilson, S., Walker, P., & Clancy-Menchetti, J. (2013). Evaluating the components of an emergent literacy intervention for preschool children at risk for reading difficulties. *Journal of Experimental Child Psychology*, 114, 111–130. https://doi.org/10.1016/j.jecp.2012.08.010

Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly*, 13(2), 263–290. https://doi.org/10.1016/s0885-2006(99)80038-6

Maul, A., & Ambler, K. (2014). Embedding language therapy in dialogic reading to teach morphologic structures to children with language disorders. *Communication Disorders Quarterly*, 35(4), 237–247. https://doi.org/10.1177/1525740114525657

Merga, M. K. (2017). Interactive reading opportunities beyond the early years: What educators need to consider. *Australian Journal of Education*, 61(3), 328–343. https://doi.org/10.1177/0004944117727749

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. National Institute for Literacy.

Opel, A., Ameer, S., & Abound, F. (2009). The effect of preschool dialogic reading on vocabulary among rural Bangladeshi children. *International Journal of Educational Research*, 48(1), 12–20. https://doi.org/10.1016/j.ijer.2009.02.008

Pfeiffer-Flores, E., Fernandes-Pires, L., & Barbosa Alves de Souza, C. (2014). Dialogic reading of a novel for children: Effects on text comprehension. *Paidéia*, 24(58), 243–251. https://doi.org/10.1590/1982-43272458201412

Pillinger, C., & Wood, C. (2013). A small-scale comparison of the relative impact of dialogic and shared book reading with an adult male on boys' literacy skills. *Journal of Early Childhood Literacy*, 13(4), 555–572. https://doi.org/10.1177/1468798413491975

Pillinger, C., & Wood, C. (2014). Pilot study evaluating the impact of dialogic reading and shared reading at transition to primary school: Early literacy skills and parental attitudes. *Literacy*, 48(3), 155–163. https://doi.org/10.1111/lit.12018

Pluye, P., Robert, E., Cargo, M., Bartlett, G., O'Cathain, A., Griffiths, F., Boardman, F., Gagnon, M., & Rousseau, M. (2011). Proposal: A mixed methods appraisal tool for systematic mixed studies reviews. Retrieved May 25, 2019, from http://mixedmethodsappraisaltoolpublic.pbworks.com

Rahn, N., Grygas-Coogle, C., & 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. https://doi.org/10.1177/0022466915622202

Reese, E., Leyva, D., Sparks, A., & Grolnick, W. (2010). Maternal elaborative reminiscing increases low-income children's narrative skills relative to dialogic reading. *Early Education and Development*, 21(3), 318–342. https://doi.org/10.1080/10409289.2010.481552

Roberts, T. (2008). Home storybook reading in primary or second language with preschool children: Evidence of equal effectiveness for second-language vocabulary acquisition. *Reading Research Quarterly*, 43(2), 103–130. https://doi.org/10.1598/RRQ.43.2.1

Saracho, O., & Spodek, B. (2010). Parents and children engaging in storybook reading. *Early Child Development and Care*, 180(10), 1379–1389. https://doi.org/10.1080/03004430903135605

Schickedanz, J., & McGee, L. (2010). The NELP report on shared reading interventions (chapter 4): Extending the story. *Educational Researcher*, 39(4), 323–329. https://www.jstor.org/stable/27764604. https://doi.org/10.3102/0013189X10370206

Scholastic Kids & Family Reading Report. (2018). The rise of read-aloud. Retrieved June 6, 2022, from https://www.scholastic.com/readingreport/rise-of-read-aloud.html

Senechal, M. (2006). Testing the home literacy model: Parent involvement in kindergarten is differently related to grade 4 reading comprehension, fluency and reading for pleasure. *Scientific Studies of Reading*, 10(1), 59–87. https://doi.org/10.1207/s1532799xssr1001_4

Senechal, M., & LeFevre, J. (2001). Storybook reading and parent teaching: links to language and literacy development. In P. Britto & J. Brooks-Gunn (Eds.), *The role of family literacy environments in promoting children's emerging literacy skills. New directions for child and adolescent development*. Jossey-Bass Inc.

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 Child Development and Care*, 184(11), 1531–1549. https://doi.org/10.1080/03004430.2013.862532

Simsek, Z., & Erdogan, N. (2015). Effects of the dialogic and traditional reading techniques on children's language development. *Procedia - Social and Behavioral Sciences*, 197, 754–758. https://doi.org/10.1016/j.sbspro.2015.07.172

Smetana, L. (2005). Collaborative storybook reading: Bring caregivers and at-risk kindergarten students together. *Reading Horizons*, 45, 283–320. https://scholarworks.wmich.edu/reading_horizons/vol45/iss4/2

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

Towsen, J., & Gallagher, P. (2014). Training head start parents in dialogic reading to improve outcomes for children. *International Journal of Child Health Human Development*, 7(3), 287–296. https://www.researchgate.net/

publication/326552523_Training_Head_Start_parents_in_dialogic_reading_to_improve_outcomes_for_children

Towsen, J., Gallagher, P., & Bingham, G. (2016). Dialogic reading: Language and preliteracy outcomes for young children with disabilities. *Journal of Early Intervention*, 38(4), 230–246. https://doi.org/10.1177/1053815116668643

Tysbina, I., & Eriks-Brophy, A. (2010). Bilingual dialogic book-reading intervention for pre-schoolers with slow expressive vocabulary development. *Journal of Communication Disorders*, 43, 538–556. https://doi.org/10.1016/j.jcomdis.2010.05.006

Urbani, J. M. (2020). Dialogic reading: Implementing an evidence-based practice in complex classrooms. *Teaching Exceptional Children*, 52(6), 392–402. https://doi.org/10.1177/0040059920917694

Valdez-Menchaca, M. C., & Whitehurst, G. J. (1992). Accelerating language development through picture book reading: A systematic extension to Mexican day care. *Developmental Psychology*, 28(6), 1106–1114. https://doi.org/10.1037/0012-1649.28.6.1106

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*, 56(8), 865–873. https://doi.org/10.1111/jcpp.12352

What Works Clearinghouse. (2007). Dialogic reading: Early childhood education. Retrieved March 3, 2021, from http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=135

What Works Clearinghouse. (2010). Early childhood education interventions for children with disabilities: Dialogic reading. Retrieved March 3, 2021, from http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=136

Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, 30(5), 679–689. https://doi.org/10.1037/0012-1649.30.5.679

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, 24(4), 552–559. https://doi.org/10.1037/0012-1649.24.4.552

Zevenbergen, A., Whitehurst, G., & Zevenbergen, J. (2003). Effects of a shared-reading intervention on the inclusion of evaluative devices in narratives of children from low-income families. *Journal of Applied Developmental Psychology*, 24, 1–15. https://doi.org/10.1016/S0193-3973(03)00021-2

Claire Pillinger is a research assistant at the Centre for Teachers and Teaching Research at the University College London Institute of Education. Claire has 12 years' experience conducting research in schools. She is particularly interested in the home literacy environment, social justice in education and reducing educational inequalities. She is currently completing a PhD by publication on dialogic reading practices.

Received 28 September 2020; revised version received 21 June 2022.

Address for correspondence: Claire Pillinger, Centre for Teachers and Teaching Research, UCL Institute of Education, 20 Bedford Way, London WC1H 0AL, UK. Email: c.pillinger@ucl.ac.uk