

**Effects of training parents in Dialogic Book-Sharing: The Early-years Provision in
Children's Centers (EPICC) Study**

Lynne Murray^{a*}, Susie Jennings^b, Hayley Perry^b, Martin Andrews^a, Katherine De Wilde^a, Amber Newell^a, Alicia Mortimer^a, Emily Phillips^c, Xiaonan Liu^b, Claire Hughes^d, Edward Melhuish^e, Leonardo De Pascalis^{f,g}, Corinne Dishington^h, John Duncan^c, Peter J Cooper^a

^aSchool of Psychology and Clinical Language Sciences, University of Reading, RG6 6AL, UK

^bThames Valley Clinical Trials Unit, University of Reading,

^cMedical Research Council Cognition and Brain Science Unit, University of Cambridge

^dDepartment of Psychology, University of Cambridge

^eDepartment of Education, University of Oxford

^fDepartment of Psychological Science, University of Liverpool

^gDepartment of Psychology, University of Bologna

^hUnder Fives Service, Reading Borough Council, Reading

*Corresponding author lynne.murray@reading.ac.uk

Declaration of interest: None

Acknowledgements

We thank the following: Reading children's centers' staff for facilitating the intervention's implementation; student volunteers for data coding and management; Jack Dainty for work on randomization, Duncan Astle for data analysis; Joni Holmes, Kathy Sylva, Anna-Louise Van Der Merwe and David Wellsted for advice and guidance.

Funding: The study was funded by the Nuffield Foundation, Reference EYP/42798, and the Medical Research Council (UK) through intramural funding to the MRC Cognition and Brain Sciences Unit, University of Cambridge. Neither funder had any involvement in the study design, collection, analysis and interpretation of data, writing of the report, and decision to submit the article for publication.

Abstract

Shared picture-book reading is well-recognized as beneficial for children's early language development, especially where 'dialogic book-sharing' techniques are used. Possible benefits of dialogic book-sharing to other aspects of child development have been little investigated, and it has not been widely studied in European populations. We conducted a randomized trial of dialogic book-sharing in Children's Centers in the UK, with parents of 2-4-year-old children, hypothesizing that it would benefit parenting and a range of child developmental outcomes. Intervention group parents (n = 110) received seven, weekly, group training sessions and control parents (n = 108) the usual center input. Parenting and a range of child outcomes (language, attention, executive function, social development and emotional-behavior difficulties) were assessed on three occasions: before, after, and 4-6 month following intervention. For all study outcomes we compared controls with each of the Intention-to-Treat population and the per-protocol population (participants attending the requisite number of sessions); and, for primary child outcomes only, the population of parents who engaged well with the intervention. There were substantial benefits of dialogic book-sharing training to parental behavior during book-sharing, especially for sensitivity and cognitive scaffolding. For all three sets of comparisons there were small-medium effects of on child expressive language, and, for the per protocol and engaged populations, similar sized effects on child receptive language and attention. There was no evidence of benefit of dialogic book-sharing for the other areas of child development; we suggest that specific intervention components need to be added to standard dialogic book-sharing to effect change in these areas.

Keywords: dialogic reading, book sharing, child development, parenting intervention, parent-child interaction

Highlights

- Parental training in ‘dialogic book-sharing’ is deliverable in UK Children’s Centers
- Training led to large gains in parent book-sharing skills with their 2- to 4-year-olds
- Child language and attention showed some benefit of the intervention.
- There was no evidence of benefit to other areas of child development.

Core skills affecting pre-school children's school readiness (e.g., language, attention, managing behavior and emotions, and social relationships (Hughes et al., 2015)) show marked disparities in the UK that are linked to socio-economic inequity and associated variation in the home learning environment (DfE, 2014; Marmot et al., 2020; Melhuish et al., 2008; Ryan et al., 2006; Sammons et al., 2015; Sylva, 2014), as is the case more widely (ASPE, 2014; Perry et al., 2018). These early childhood differences commonly persist and influence later academic achievements, employment, and adjustment, thereby perpetuating inter-generational cycles of disadvantage (Allen, 2011; Centre for Social Justice, 2014). An important aspect of the home learning environment associated with child development is parental education (Hoff, 2013). Indeed, even within low-income families, the speed of infant language acquisition is positively associated with the level of maternal education (Justice et al., 2020). One key way in which parental education exerts its impact is via the amount that parents read to their child. The benefits of being read to are apparent from infancy (Leech et al., 2022), and are especially evident in the preschool years (Fletcher & Reese, 2005; Rodriguez et al., 2009). Indeed, one longitudinal study examining development from 3 to 29 years of age showed that, over and above associations with maternal education, variability in the extent to which parents read to their preschool children predicted educational attainment in adulthood (Gottfried et al., 2015).

One way in which parents reading to children can enhance child development is by providing a 'lexical reservoir' that widens child vocabulary (Crain-Thoreson et al., 2001; Hoff-Ginsberg, 1991), a key predictor of later educational progress (Morgan et al., 2015). In a recent study, Logan et al. (2019) used data from 60 popular American children's books to estimate that a daily diet of just one-book-a-day of shared reading enables children to hear 78,000 words per year. Based on this impressive tally, these authors concluded that, compared with children whose

caregivers do not read to them (estimated at 23% of the population in the USA (Khan et al., 2017)), children from literacy-rich homes who regularly enjoy sessions with multiple storybooks hear 1.4 million more words across the five years before kindergarten entry.

Aside from how *much* parents read to their child, the *quality* of parent-child interaction while reading is also important. One style of reading, known as ‘dialogic book-sharing’ (Whitehurst & Lonigan, 1998), has been shown to be particularly helpful to child language development (Mol et al., 2008). Dialogic book-sharing is fundamentally collaborative, with children encouraged to participate actively in the book-sharing process. The adult uses specific conversational techniques, such as building on the child’s interest and vocalisations, and asking questions aligned with the child’s ‘zone of proximal development’ (Vygotsky, 1978) that link the book content to the child’s own experience. These features of parental talk, as they arise in general conversational contexts, have long been known to promote infant and child language development (Fernald & Weisleder, 2011; Snow, & Ferguson, 1977; Taumoepeau, 2016), but they are particularly likely to occur in book-sharing interactions (Hoff-Ginsberg, 1991; Adrian et al., 2005; Ninio, 1983; Noble et al., 2018; Salo et al., 2016). This is especially the case when parents use picture books or text-light books where, rather than relying on a pre-specified text, they instead construct their own account of the content and adjust it to their child (Sénéchal et al., 1995). Aside from these linguistic features, picture book sharing also affords a number of specific interactive behaviors such as gaze following, pointing, and gestural and vocal animation that are associated with enhanced child cognitive processing and language (Murray et al., 2022). In an early review of 16 dialogic book-sharing intervention studies, Mol et al. (2008) reported that this method produced larger gains in children’s expressive language than ‘normal reading’

($d = .59$) (although this effect was clearer for families with 2- to 3-year-olds versus 4- to 5-year-olds or children at risk for language and literacy impairments (see also McGinty et al., 2012)).

Importantly, parents experiencing socio-economic disadvantage read less to their children than do others and, when they do read to their children, they are less likely to use ‘dialogic’ techniques (Bus & van IJzendoorn, 1995; Fletcher & Reese, 2005; Heath, 1982; Raikes et al., 2006). Interventions to foster the use of dialogic book-sharing with young children in families experiencing socio-economic disadvantage could therefore help narrow educational gaps related to family background. Indeed, a recent meta-analysis of 19 RCTs that focused on providing training in this method to parents largely living in disadvantaged conditions has shown significant benefits for child expressive and receptive language, with average effect sizes (d) of 0.41 and 0.26 respectively (Dowdall et al., 2020). Two of the studies reviewed by Dowdall et al. were conducted in poor communities in South Africa, where a six to eight-week book-sharing training program brought about substantial improvements in both child language and attention (Cooper et al., 2014; Vally et al., 2015). Notably, these benefits were mediated by gains in parental sensitivity and reciprocity whilst sharing picture books (Murray et al., 2016).

While book-sharing studies have focused principally on benefits to child language, there are grounds for hypothesizing wider benefits of dialogic book-sharing for child development (as indicated by the gains in child attention noted in the two South African studies). Several aspects of the process underpin this hypothesis. First, in dialogic book-sharing parents perform a *scaffolding* function for child cognition; that is, they help structure the child’s mental processing of the book content, responding to, as well as guiding the child’s attention to critical elements and helping them link these to the broader book content. Such scaffolding is known to support the development of the general skill of attentional structuring, promoting across-the-board

cognitive proficiency. Thus, parental scaffolding is correlated with a range of measures of child cognition, including tests of executive function, a broad range of ‘fluid intelligence’ skills concerned with regulating cognition (e.g., working memory, shifting, inhibition) (Fay-Stammbach et al., 2014; Hughes & Devine, 2019; Lengua et al., 2014), and verbal and performance IQ (Obradovic et al., 2016), as well as with child classroom behavior (Neitzel & Stright, 2003). Although a parent may help to structure their child’s mental activity in *any* complex setting, book-sharing provides particularly good opportunities for joint attention and cognitively enriching parental support (Murray et al., 2022). Indeed, in a low-income Pakistani community, scaffolding behaviors during book-sharing were found to predict a wide range of child cognitive skills (Obradovic et al., 2016).

A second important aspect of dialogic book-sharing relevant to child development is that it is a favored context for parents to engage in talk about mental states and to use specific structural characteristics of language, such as complement clause constructions (Adrian et al., 2007; Brandt et al., 2016; Noble et al., 2018). Each of these aspects of parental talk has been shown to predict children’s later theory of mind (ToM). This capacity is central to children’s social functioning, and entails understanding that others may have different mental experiences (i.e., feelings, beliefs, perspectives) from one’s own (e.g., Brandt et al., 2016; Boeg et al., 2021; Devine & Hughes, 2019; Ensor et al., 2014). Notably, mental state talk is also known to predict children’s executive function (Baptista et al., 2017). Furthermore, both well-developed child theory of mind and executive function are themselves associated with lowered rates of behavior problems (e.g., Hughes et al., 2020; Hughes & Ensor, 2007; Roman et al., 2016), as is the child language proficiency associated with dialogic book-sharing (Morgan et al., 2015; Petersen et al., 2013). There are, therefore, evidential grounds for hypothesizing that, through the cognitive

mediators of enhanced social understanding, executive function, and language skills, the beneficial effects of dialogic book-sharing on caregiver-child interactions may also extend to reducing children's behavioral problems.

Aside from any secondary benefits of dialogic book-sharing to child behavior problems achieved by enhancing performance in other areas of development, it can also be hypothesized that its core features may confer *direct* benefits to child behavior problems, since they comprise many of the 'positive parenting' qualities that have been found to be associated with low rates of these child difficulties. These parenting qualities include the provision of consistent routines, in this case book-sharing, that offer child-centered time, and warm, sensitive interactions that are developmentally appropriate and that follow the child's lead, using praise and encouragement, rather than coercive or critical practices (Chacko et al., 2018).

Importantly, intervention research highlights the malleability of each of early executive function, social understanding, and child behavioral problems, including when using book-sharing as a vehicle for intervention delivery. Thus, a recent meta-analysis of 35 RCTs of interventions to support executive function in preschool samples showed positive effects on a wide range of skills (Pauli-Pott et al., 2021). Of particular relevance to the current study, Howard et al. (2016) have shown that book-sharing with preschool children that includes explicit executive function-relevant components leads to sustained improvement in these skills (i.e., working memory and shifting), with effects apparent after as little as three weeks. With regard to social understanding, a meta-analytic review of 45 studies of theory of mind training for children (mean age five years) showed a substantial overall benefit (Hofmann et al., 2016). Similarly, a more recent review of such training, including 10 studies with preschool children, showed a large effect size (2.51) for this age group. Notably, these latter studies generally included use of

picture books accompanied by specific interrogative input relevant to the mental states of book characters (Roheger et al., 2022). Finally, a meta-meta-analysis of 26 meta-analyses of parenting training programs for child conduct problems that included several hundred pre-school and school age children showed benefits to both observed and parent-reported child difficulties (Mingebach et al., 2018). As for executive function and social understanding, intervening with parents to address child behavior problems has also included using dialogic book-sharing. Thus, Chacko et al. (2018) showed that a core dialogic book-sharing program, with integrated parent training elements (e.g., using ‘labeled praise’, and ‘ignoring and distraction’) was of significant benefit to child behavior.

The current study

Despite the accumulating evidence for the benefits of dialogic book-sharing in a range of populations, to date there has been little research on this topic in the UK and the rest of Europe. This is an important gap, particularly in the UK context, where socio-economic inequities have shown a marked increase in the last decade, and where there is therefore a pressing need for interventions to address their impact on child development (Marmot et al., 2020). One recent UK study (Burgoyne et al., 2018) did evaluate an intervention that included dialogic book-sharing, with positive results; however, as this intervention was multifaceted, the contribution of the book-sharing element is unclear. We therefore conducted an RCT of a dialogic book-sharing intervention with parents of preschool children in the UK. Further, given the case outlined above for the possible wide-ranging benefits to child development of dialogic book-sharing interventions, we built on previous research by extending examination of the impact beyond child language. Thus, we examined whether parent training in book-sharing affected children’s

wider cognitive functioning, as well as their social and behavioral-emotional development. We predicted that improvements in parenting would mediate any benefits of the intervention to child development, and accordingly we assessed parent child-interactions. We also examined the potential moderating effects of key child and parent factors known to affect child development and parenting (e.g., child age, and parental education, mental health (depression, anxiety), and parenting stress (Cicchetti et al., 1998; Murray et al., 2020; Robinson & Emde, 2004; Weissman, et al., 2006; Weissman, Pilowsky, et al., 2006). Finally given that participant commitment to interventions is important to their effectiveness (e.g., Garvey et al., 2006; Gennetian et al., 2019; Nix et al., 2009; Piotrowska et al., 2017; Waanders et al., 2007), including in the context of interventions to promote use of shared reading with young children (de Bondt et al., 2020), we also examined parental attendance and engagement in the dialogic book-sharing program as predictors of outcome.

Ethical approval for the trial was obtained from the University of Reading Research Ethics Committee (reference number: N15/09/084), the trial was registered in March 2017 (ISRCTN 28513611), and a protocol paper was published (Murray et al., 2018).

Method

Context

The ‘Impact of Early Years Provision in Children’s Centers (EPICC)’ study was based in Children’s Centers in the city of Reading, UK. The population of Reading is broadly representative of the general UK population, including multi-ethnic communities and significant areas of deprivation. In Reading, over 3,000 children under 5 years live within the 20% most deprived areas (i.e., Lower Super Output Areas (LSOAs) [National IDACI index]), and most Children’s Centers are situated in these deprived areas. Children’s Centers are publicly funded

and are free to users. They aim to reduce inequalities between families experiencing disadvantage and others by offering support for family functioning and parenting, as well as child health and development from birth to the start of school. To avoid stigmatizing users, Children's Centers are inclusive, rather than restricted to families regarded as the 'most needy', and those attending are broadly representative of the locations where they are based (DfE 2011; Sylva, 2014). The support provided for child development by Children's Centers is regulated by the Office for Standards in Education, and is delivered following the national curriculum, the Early Years Foundation Stage (EYFS; <https://www.gov.uk/government/publications/early-years-foundation-stage-framework--2>). The EYFS adopts a wholistic approach to ensuring that the major areas of child development are covered in a play-based curriculum. These areas include communication and language, and social and emotional development, with book time, or story time, usually comprising part of routine support for communication and language. Research into the benefits of Children's Centers has shown substantial continuity of effects of family and child characteristics assessed at intake on child development, but also evidence for parents' use of Children's Centers helping to ameliorate, although not overcome, the impact of financial disadvantage on some family and maternal outcomes and, to a lesser extent, child social, although not cognitive, development (Sammons et al., 2015).

Aside from the provision from Children's Centers, an additional point to note regarding the context of the current study is that there is a national BookStart program in the UK that involves providing all children with two books in their first year, and a further book at age three-to-four years (booktrust.org.uk). A recent meta-analytic review of programs providing free books to families, including the UK BookStart program, has shown some benefits to the home learning

environment, generally measured in terms of the frequency of parent reading to the child, although effects on child language have not been studied in UK samples (de Bondt, et al., 2020).

Randomization of Children's Centers

Inspection of the populations (electoral wards) in which the Reading Children's Centers were located showed one to be an outlier in terms of the Index of Multiple Deprivation (IMD) and ethnic group, and this center was excluded from the study. The remaining 12 centers were randomly assigned by an independent statistician either to the intervention condition (6 centers) where parents received training in book-sharing, or to the control condition (6 centers). The first two centers were allocated by simple randomization using random number generation in Excel. The remaining centers were then allocated by minimization, a method of adaptive stratified sampling. Allocation was based on minimizing the (mean) differences in IMD and ethnicity between the intervention and the control group (Pocock, & Simon, 1975). This method was used as we had a small number of centers to randomize, and there were several predictive factors we wished to take into account. Given these constraints, minimization is preferred over stratified randomization, and leads to greater balance regarding the predictive factors (Scott et al., 2002; Treasure & MacRae, 1998).

Sample size and recruitment

The required study sample size was based on a meta-analysis of book-sharing interventions (Dowdall, 2015), where three trials (two in the US and one in South Africa) using a similar format to that used in the current study showed an average effect size on language development, our key primary outcome, to be 0.88. We therefore planned for a mid-range medium effect size. With an effect size of $d = 0.66$, within the cluster design, an intervention group and control group sample of 96 each were required (with a two-sided test, $\alpha = 0.05$

and $\beta = 0.90$, intra-cluster correlation = 0.04). With an addition of 10% participants to account for sample loss, a total sample of 214 was required – i.e., two groups of 107.

All parents (or other carers such as grandparents) attending the 12 study Children's Centers with a child aged between 28 and 45 months old, and who regularly spoke English at home, were invited by the trial manager to participate in our study. Children with a significant developmental disorder (e.g., autism, Down's syndrome) were excluded. All participants were told that the study concerned investigating the impact of Children's Center provision on families and child development and, for participants attending intervention group centers, the nature of the dialogic book-sharing program was described. All study participants were told they would receive gratuities for assessment sessions (amounting to a total of £75 or US\$100).

Of 343 parents who were contacted about the study, 96 (28%) either declined to participate or did not respond, and an additional 28 (8%) of those initially approached were found to be ineligible. Numbers of decliners/non-responders were 61 out of 190 (32%) from intervention centers, and 35 out of 153 (23%) from control centers. 108 participants were recruited into the Control group and 110 into the Intervention group. (One additional participant originally assigned to the Intervention group was subsequently removed from analyses, due to a diagnosis made by an independent clinician during the study that made them ineligible). As per trial guidelines for good practice, we present the progress of the sample through the study in the CONSORT figure ([Consort - Welcome to the CONSORT Website \(consort-statement.org\)](http://www.consort-statement.org)) (see Fig. 1).

Procedure

The study was coordinated from the School of Psychology and Clinical Language Sciences of the University of Reading, with intervention sessions run within Children's Centers.

Assessments of child development and parenting were carried out at baseline (i.e., before the intervention phase), immediately post-intervention, and at follow-up approximately 4 – 6 months after the end of the intervention. Baseline and post-intervention assessments were conducted in Children’s Centers and the follow-up assessment within University of Reading premises.

Assessments were conducted by two experienced trained psychologists, each with a Master’s degree in child development. Each assessment session lasted, in total, between 90-120 minutes and were scheduled to suit the family’s routine. Many of the child assessments took the form of games and puzzles, and included toy play, and book-sharing with the parent. Breaks were taken for child refreshment, timed to ensure maximal child engagement, as judged by assessor and parent. If children were judged to be fatigued before completion, a second assessment session was scheduled. Assessors were blind to group, and they had no contact with staff delivering the intervention. Participants and Children’s Center staff were directed not to discuss group assignment or any aspect of the study with the assessors.

The intervention

The book-sharing intervention was an adaptation of one originally developed in the US by Whitehurst et al. (1988) and shown to be of benefit in numerous studies, principally conducted in the US, to child language development (Dowdall et al., 2020). The version we used was modelled on one previously delivered in South Africa, being delivered to small groups of parents, weekly, over a period of seven weeks (Cooper et al., 2014; Murray et al., 2016; Vally et al., 2015). This duration is at the upper end of the range of effective interventions in the Dowdall et al. (2020) review. The core principles are that parents are trained in how to support their child’s interest and active engagement, rather than simply ‘reading’ to their child.

Responsiveness is emphasized that is sensitive to the child’s developmental capacity and

experience, as well as the importance of a positive encouraging approach. The intervention was delivered in Children's Centers by two trained research facilitators with a background of working in the early years context. They received weekly supervision from LM and PJC. During these supervisions the facilitators discussed the individual participants' engagement with the program, including any difficulties they were experiencing and, where these were present, strategies were devised for dealing with them. In accordance with the literature (e.g., Nix et al., 2009), the supervision sessions helped ensure consistency of facilitators' ratings of parent engagement (see details below in Measures). The weekly, facilitator-led small-group sessions were conducted with groups of up to six parents, during which children were cared for in an adjoining play space by Center staff. Group sessions lasted approximately 50 minutes and were followed by a brief period of one-to-one discussion between the facilitator and each parent (together with their child). A meta-analysis of effects of book-sharing interventions showed a significant effect of the amount of training parents received (Dowdall et al., 2020), and we judged that attendance at five sessions constituted the minimum adequate number for our intervention.

The intervention provided parents with guidance on how to share books with their child in ways that best support child development, based on research on the parenting predictors of child cognitive and socio-emotional functioning (e.g., use of cognitively enriching extensions to child utterances to promote language development (Taumoepeau, 2016), and of mental state talk to promote social understanding (see Devine & Hughes, 2018 for a meta-analytic review)). Facilitators delivered the intervention to parent groups via Powerpoint presentations that included demonstration video examples of parents using good book-sharing practices with children (see Figure S1 for example Powerpoint slides). The content of each intervention session

was specified in a facilitator user manual containing copies of the materials shown to parents, along with related instructions on delivery. Sessions first recapitulated the core learning points of previous ones, and then focused on a particular theme, using a designated children's book, ('the book of the week'), covering specific book-sharing techniques to enhance child development (see Table 1 for session books and content). For example, in the second session guidance was given on responding to children by building on what they say and making links between the book content and the child's own experience; and in the fourth session, guidance was provided on highlighting the book characters' emotional expressions and experiences. During the group sessions, the facilitator invited parents to comment on the videos and to join in discussion about the recommended techniques. Sample illustrations from the week's book were included in the Powerpoint presentations and used to facilitate group discussion about how the techniques covered in the session could be applied when sharing the book with the child. The books used during training were selected to match each session theme and to provide opportunities to rehearse the sessions' techniques. For example, the fourth session concerning emotions used 'Hug', by Jez Alborough, a book about a baby monkey being separated and then reunited with its parent; and the seventh session concerning relationships used 'The wrong side of the bed', by Edward Ardizzone, a book about the occurrence and resolution of an everyday disagreement between a child and his parents. At the end of each group session, in order to generalize skills across different books, the application of the DBS techniques covered that week were discussed in relation to the previous weeks' books.

Books were either wordless, or text-light, since this format has been shown to elicit mutually responsive parent-child interactions, and to afford elaborated dialogic talk, including

talk concerning the mental states of book-characters (Noble, et al., 2018; Peskin & Astington, 2004; Sénéchal et al., 1995).

Following the group session, the facilitator met individually with each parent and their child for five to ten minutes, as required, and invited them to share the book of the week together, using the techniques covered in the group session. This provided an opportunity for the facilitator to encourage and support parents, as well as to model particular book-sharing techniques if parents were uncertain about them. Any challenges parents experienced in implementing book-sharing at home were discussed, and strategies for managing these explored. As noted, sessions recapitulated key points from the previous ones, thereby reducing the possible impact on participants of their having missed any given session. In addition, to further address this issue, if parents had missed a session, the facilitator made sure the relevant learning points were well-understood in the subsequent one-to-one meeting.

Each book of the week was given to participants to take home with them at the end of the session, along with a card containing brief reminders of the session's main points (see Figure S2 for the card used for session 1). Participants were encouraged to practice sharing the book with their child regularly over the coming week, and they reported back about their experiences at each group session. At approximately two months and three-to-four months after the intervention, parents were sent an additional book, as a booster (see Table 1), with accompanying suggestions on how it might be used with the child.

To ensure fidelity of delivery of the intervention, in addition to its being precisely specified in the manual and session Powerpoint presentations, two of the PIs (LM, PJC) sat in on a random sample of training sessions of both facilitators and found them on all occasions to be adhering strictly to the specified training program.

Measures

Sociodemographic and Parent Characteristics

Parents completed standard proformas to record information relevant to analyses of outcome. These included baseline measures of child age, sex, birth order and presence of siblings, and parental income, education, multilingualism, and amount of reading to the child prior to the study (measured at follow-up to prevent priming bias in control group participants (Rodrigues et al., 2015)- see Supplementary materials S3 for details, including measures taken to ensure reliability, and validity data).

Parents also reported, at each assessment point, on their own mental state, using the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), and on parenting stress, using the Parenting Stress Index (Abidin, 1995).

Child Assessments

Child functioning was assessed in four developmental areas, namely, cognitive development (the primary outcome), social development, behavior difficulties, and emotion regulation. In line with common, recommended, practice for RCTs (Freemantle et al., 2003; ICH Harmonised Tripartite Guideline, 1999; Song et al., 2013), individual component assessments were administered and robust composite measures derived to reflect the underlying core developmental constructs (Landis et al., 2000). This was done for key dimensions of the primary outcome of cognitive development, and for the key secondary outcome of social development, in line with the procedures specified in the pre-published study protocol (Murray et al., 2018). This method avoids reliance on single component measures that might be subject to some methodological or conceptual limitation, and is a strategy that is useful in increasing statistical power (Song et al., 2013). The assessment measures and associated outcome variables are listed below. They included parent reports and direct child assessments. Assessments were selected

according to child age, with the most comprehensive set being conducted at follow-up, the main outcome point. (See below and Table 3 regarding those measures not used at all three time points). Those delivered by tablet or using standard proformas were scored during administration. Observational assessments were videotaped and subsequently scored by trained independent researchers. Random samples of videotapes for each observational measure were rated by a second researcher (see Results section for reliability values).

Cognitive Development (Primary outcome). We assessed three dimensions of cognitive development: language, attention, and executive function.

Language. Two measures of expressive language were administered (a) the vocabulary sub-test of the Early Years Toolbox (EYT; Howard & Melhuish, 2017) (score = total number of correct words); (b) parent report, using the Communicative Development Inventories (CDI; Fenson et al., 2000) to record child words spoken (score = total words). Two measures of language comprehension were administered (a) the Clinical Evaluation of Language Fundamentals (CELF-2; Wiig et al., 2004), using the subscales “sentence structure”, “concepts and following directions”, and “basic concepts” (score = mean percentage correct of the three subscales) (follow-up only); (b) parent report, using the CDI to record child words understood (score = total words).

Attention. We produced a composite measure of attention, comprising (a) sustained attention, using the Early Child Vigilance Task (ECVT; Goldman et al., 2004; Vally et al., 2015) (score = the percent time attending to a video cartoon stimulus over seven minutes); (b) persistence of attention, using the Three Toy Play task (Cooper et al., 2014; Kannass et al., 2006) (score = longest time attending to one of three available toys); (c) quality of attention, using the Three Toy Play task (score = mean of time-sampled 4-point scale ratings); (d) consistency of

attention, using the EYT Go-no-go task (Howard & Melhuish, 2017) (score = standard deviation of time to respond in correct ‘go’ trials); (e) parent report, using the Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001; items 15 and 25), and the EYT Children’s Self-regulation and Behavior Questionnaire (CSBQ; Howard & Melhuish, 2017; item 4) (score = item total). (See Statistical Methods below for computation of composite measure).

Executive Function (EF). We produced a composite measure of executive function comprising (a) general EF, using the Block Design subtest from the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 2012) (score = total points) (post and follow-up only); (b) shifting, using the EYT Card Sort task (score = total correct) (follow-up only); (c) working memory, using the Following Instructions task (Jaroslawska et al., 2014), where the child follows instructions to perform actions on different types of objects of different colours (score = total elements correct (actions, objects, colors) (post and follow-up only); (d) working memory, using the Digit Span task (Alloway, 2007), where the child repeats back digits in strings of increasing length (score = total trials correct) (post and follow-up only); (e) impulse control, using the EYT Go-no-go task (score = proportion of correct no-go trials minus the proportion of incorrect go trials); (f) on-task persistence quality, using the Frustration task of the Laboratory Temperament Assessment Battery (Lab-TAB; Goldsmith & Rothbart, 1993), where the child was given the wrong keys to open a perspex box containing an attractive toy (score = mean of time-sampled 5-point scale ratings) (baseline and follow-up only); (g) parent report, using the EYT CSBQ behavioral and cognitive regulation subscales (score = total of the two subscales). (See Statistical Methods below for computation of composite measure).

Social Development (Secondary outcome). We produced a composite measure of social development comprising (a) theory of mind (ToM), using a battery of nine tasks, each of which

required the child to understand that others' perspective, desires, knowledge or beliefs, differ from their own; the tasks comprised Penny Hiding (Baron-Cohen, 1992), Diverse Desires 1 (Repacholi & Gopnik, 1997), Diverse Desires 2 (Wellman & Liu, 2004), Diverse Beliefs (Wellman & Liu, 2004), Knowledge Access (Wellman & Liu, 2004), Recall of False Belief (Hughes et al., 2011), Explanation of False Belief (Hughes et al., 2011), Explicit False Belief (Wimmer & Perner, 1983), Second Order False Belief (Wimmer & Perner, 1983). (All ToM tasks were scored as pass/fail, apart from Penny Hiding which itself comprised a set of pass/fail scores, and was scaled to 0-1), and the summary ToM score was the total for all tasks (i.e., 0-9)); (b) expressive emotion understanding, using the Puppet task (Denham, 1986), where the child had to name a puppet's different facial expressions of emotion (score = total emotions correctly labelled); (c) empathy, using the Help task (Murray et al., 2016), where the researcher enacted losing a glue stick they needed (score = extent x swiftness of child help given); (d) altruism, using the Dictator Game Altruism task (Benenson et al., 2007; Decety, 2015), where the child had to apportion sweets to two boxes, one for themselves, and the other for another child (score = number of sweets allocated to other child) (follow-up only); (e) parent report, using the prosocial scale of the SDQ (score = total). (See Statistical Methods for computation of composite measure).

Behavior Difficulties (Secondary outcome). Two measures were used: (a) Defiance, using the Don't Touch task (Kochanska & Askan, 1995; NICHD, 1998), where the child was instructed not to touch three attractive toys (score = mean of time sampled 6-point scale ratings) (baseline and follow-up only); (b) parent report, using the 'emotional' 'hyperactivity', 'conduct' and 'peer problems' subscales of the SDQ (Goodman, 2001; score = total of subscales).

Emotion Regulation (Secondary outcome). Two measures were used: (a) reactivity using the Frustration task of the Lab-TAB (Goldsmith & Rothbart, 1993; score = weighted mean of 6-point time-sampled anger and distress ratings) (baseline and follow-up only); (b) parent report, using relevant items from the EYT CSBQ (i.e. 1, 6, 7, 10, 13, 14) (Howard & Melhuish, 2017; score = sum of items).

Parent-child Book-sharing Interactions (Secondary outcomes)

The quality of book-sharing was assessed at baseline and post intervention. On both occasions parents were asked to share the same text-light book with their child ('Yes' by Jez Alborough) in the way they would at home. Interactions were filmed and speech transcribed, with all transcripts checked by a second researcher. Five minutes of interaction were rated on measures of parental behavior and parent-child interaction. The measures concerned dimensions of book-sharing that the intervention was designed to enhance, and they therefore provided an objective measure of how well parents had implemented the strategies covered in the training program. Videos and transcripts were scored by trained researchers blind to group and child outcome. Random samples were scored by independent trained assessors to establish interrater reliability (see Results below for reliability values). (The rating manual is available from the first author.)

Parent Behavior

We assessed three dimensions of parent behavior.

Sensitivity. This concerns parental appropriate and warm responsiveness to the child (Cooper et al, 2014; Murray et al, 2016). Key aspects include the parent's awareness of the child's focus of interest (e.g., gaze direction, pointing, efforts to turn the page) and their communication, as well as the extent to which parental responses to these behaviors were

supportive and well timed. The level of sensitivity was rated from videos (score = rating on 5-point scale).

Cognitive Scaffolding. We used transcripts to produce a composite measure of scaffolding, comprising (a) enrichment (percentage of child utterances responded to with a cognitively enriching extension); (b) elicitation of child involvement (questions, invitations to complete an utterance); (c) attention structuring (references to book elements - actions, characters, experiences); (d) complexity (utterances specifying contrasts and causes, and sentential complements). The Cognitive Scaffolding score was the mean of z scores for the four individual measures.

Mental State Talk. This comprises references to mental states (desires, cognitions, feelings) (Devine & Hughes, 2018), and their occurrence within complex utterances (see (d) above) and was scored from transcripts (score = frequency).

Parent-child Behavior

We assessed two aspects of parent-child reciprocity.

General Reciprocity. This concerns shared affect (e.g., smiles, expressions of surprise, concern) and joint attention to the book (e.g., gazing and pointing together at the same part of the page), vocal exchanges and gestural turn taking (e.g., stroking motions on depicted book characters), and mutual gaze (Cooper et al., 2014; Murray et al., 2016), and was rated from the videos (score = rating on 5-point scale).

Verbal Reciprocity. This was event rated from transcripts (score = average of the % of child utterances responded to by the parent, and the % of parent utterances responded to by the child).

Intervention Group Measures

Attendance

Facilitators recorded participants' attendance at weekly sessions in order to identify those completing the requisite number of sessions (i.e., at least five of the seven sessions) for subsequent per-protocol analyses.

Engagement

After each session, the facilitators completed two ratings, each on a 4-point scale, concerning each parent's engagement in the session. These were based on the criteria used in the literature to measure engagement in group parenting interventions (e.g., Dumas et al., 2007; Nix et al., 2009), and established as reliable (Dumas et al., 2007). The first rating concerned the *extent* of engagement, based on parental active participation in group discussion (scored from 'not at all' to 'very much'); and the second concerned engagement *quality*, that is, the value parents placed on the training, based on their attentiveness, positive attitude, and acceptance of the concepts presented in the session (scored from 'not at all' to 'highly'). These weekly ratings on the two scales were used to derive an overall measure of engagement at the end of the training program, with a mean score ≥ 2.5 defining an engaged subgroup.

Implementation and Acceptability

Participants completed a brief questionnaire after each session. To measure their implementation of the program, they were asked how many days they had done book-sharing with their child over the previous week (score = 0-7), and how much of the time when they were book-sharing they had used the techniques covered in the previous session (rated on a four-point scale from 'not at all' to 'a lot'). To measure acceptability, parents rated how helpful they had found each session, and their enjoyment of book-sharing with their child over the previous week (each rated on a four-point scale 'not at all' to 'very much').

Statistical Analyses

Our analyses were conducted by the Thames Valley Clinical Trials Unit according to widely accepted principles outlined for randomized trials (Campbell et al., 2000; Eldridge & Kerry, 2013). Three sets of analyses were conducted. The first set were conducted according to ‘intention to treat’ (ITT) - that is, comparing the control group with all those recruited into the Intervention group, regardless of whether they had completed the requisite number of sessions. A second set of analyses compared the control group with the subgroup of intervention group participants who attended the number of sessions defined as adequate, that is, the ‘per-protocol’ intervention participants. Third, in exploratory analyses, we compared the primary child outcomes between the control group and the engaged subgroup of intervention participants.

Statistical Methods

As specified in the pre-published study protocol (Murray et al., 2018), the study was designed to have 90% power to detect a standardized effect of 0.66 for each of the primary outcome measures. Notably, while this calculation allowed for a 10% dropout rate, multiple imputation for missing data was used in several ITT analyses, leading to increased power for each primary outcome measure and the key secondary outcome of child social development. Before effects of the intervention were analyzed, scores on the different measures were examined, and some measures were found to be unsuitable for analysis due to optimum baseline scores, with no room left for improvement (i.e., maternal reports of child expressive and receptive language on the CDI, and the Lab-TAB emotional reactivity assessment). Inspection of the distributions and inter-correlations of individual components of the child composite measures (attention, executive function, and social development) confirmed their suitability for deriving composite variables. This was done using mean z scores of individual components for the

attention composite, and factor analysis for executive function and social development composites. Results for these final composite measures are reported below within the current paper; and, in line with recommendations for RCTs using composite measures (Freemantle et al., 2003), results for the individual components of composites measures are included in Tables S4 and S5.

As specified in the study protocol (Murray et al., 2018), the comparison of the intervention and control groups was carried out using linear mixed multilevel models. These included random effects for Children's Center, and fixed effects for intervention and relevant baselines. The inclusion of center as a random effect takes into account the clustered design (i.e., that we randomized the Centers to intervention, and not the participants) and the associated correlation between participants within a Children's Center (Campbell et al., 2000; Eldridge & Kerry, 2012). Child age, gender, and parent education and family multilingualism were all prespecified as co-variates and were included in the models, as well as, where it contributed independently to child primary outcomes, the amount of time parents read to their child at baseline. Where child age meant that baseline assessment was not appropriate (i.e., receptive language CELF, and some components of executive function), the expressive language score at baseline was used as a covariate. Where assessments took place at multiple time points, a longitudinal model was fitted by additionally including visit, visit by intervention interaction, and visit by baseline interaction into the specified linear mixed multilevel model. When distributions required them, transformations and other models (e.g., negative binomial) were used. Statistical tests were two-tailed and performed using a 5% significance level. No adjustments were made for multiple comparisons because the primary outcomes were associated with each other and adjustment would have over-corrected (Schulz & Grimes, 2005). Adjusted

means and 95% confidence intervals were derived from the models, and standardized effect sizes that accounted for clusters and intra-cluster correlations produced. (Effect sizes were described using Cohen's definitions to define small (.2 - .5), medium (.5 - .8) and large (>.8) effects (Cohen, 1988)). Moderation analyses, including relevant interactions with treatment and visit, were performed for primary outcomes (only within the ITT population) using linear mixed models. Moderators were parental income, education, multilingualism, mental health (HADS and PSI), child gender, age, ethnic group, birth order, presence of siblings, and baseline parenting (sensitivity measure). Mediation analyses were conducted using structural equation modelling (again, only within the ITT population). Analyses were performed using SAS 9.4, and multiple imputation and mediation using R.

Results

Sample

Table 2 shows the key demographic characteristics of the groups at baseline. In the intervention group, 101 (92%) and 97 (88%) families were seen at post and follow-up assessments respectively. Corresponding numbers for the control group were 103 (95%) and 100 (93%). Of the intervention group participants, 102 (93%) were seen for at least one of post or follow-up assessments, and of the control group the number was 105 (97%). The large majority of those recruited into the intervention group - 94 out of 110 (85%) - completed the requisite number of sessions (mean = 6.8, SD = 0.48), and formed the intervention per-protocol population. Of the remaining 16 who completed fewer than five sessions, most either did not attend at all (n = 8) or attended only one session (n =4). (Given the distribution of session attendance- that the overwhelming number of participants (106 of 110) either attended all/almost

all sessions, or else none/almost none- no analyses of continuous dose effects could be made). Of those completing the program, 69% were rated as having engaged well.

Given the random allocation of groups, no formal comparisons were made of the demographic characteristics of intervention and control group participants (Altman, 1985; Chen et al. 2020; Senn, 1994). However, it can be seen in Table 2 that, compared to the controls, within the ITT intervention group multilingualism and non-White ethnic status were rather more common, and the time parents spent reading to the child at baseline was somewhat less. Further, although the demographic characteristics of those who engaged with the intervention were broadly similar to those of control group participants, it was notable that those who did *not* engage well, and particularly those who failed to attend the requisite number of sessions (the 16 individuals who did not form part of the per protocol population) appeared to be at somewhat higher risk than those who did engage in terms of low income and education (see Table 2).

Implementation and Acceptability of Intervention

Regarding implementation, on average, participants reported sharing books with their children most days of the week throughout the program (mean days per week 4.84 ($SD=1.46$)). In addition, most reported that when they shared books with their child, they used the training techniques covered in the sessions either ‘quite a lot’ or ‘a lot’ (range across the seven sessions = 88%-99% (mean = 92.63%) of participants,).

Regarding acceptability, the great majority of participants rated the helpfulness of sessions as ‘quite a lot’ or ‘very much’ (range = 96% - 100% (mean 98.48%) of participants, across the sessions); and they similarly rated their enjoyment of book-sharing with their child as ‘quite a lot’ or ‘very much’ (range = 88% - 97% (mean 93.53 %) of participants, across sessions).

Inter-rater Reliability for Observational Measures

For child measures, reliability was as follows: attention ECVT intra-class correlation (ICC) = .86); Three Toy Play task longest bout ICC = .99, mean play quality ICC= .93; executive function Lab TAB on-task persistence ICC =.97; ToM tasks kappas range = .86-1; and Behavior difficulties Don't Touch task defiance ICC = .94. For the book-sharing interaction measures, reliability for both video-based and transcript measures was excellent, with ICCs ranging from .92 to 1.

Intervention analyses

We present the three sets of intervention analyses below, examining in turn effects for the ITT population, the per protocol population, and the engaged population. In addition to the prespecified covariates used for all measures (i.e., child age, gender, and parent education and family multilingualism), these analyses also included the amount of baseline book-sharing as a covariate for child expressive language (EYT Vocabulary), as it made a significant independent contribution to this primary child outcome ($p < .05$).

Intention to Treat (ITT) Intervention Group vs. Control Analyses

Effect sizes for comparisons between the ITT intervention group and controls at post and follow-up are shown in Table 3a. Adjusted mean scores at post and follow-up for the Primary outcomes are shown in Table 4, and those for secondary outcomes in Table S6. (Non-adjusted mean scores at baseline are shown in Table S7, and analyses of individual components of composite measures are shown in Table S4). Effects of the intervention on child development were not statistically significant. As can be seen in Table 3a, at follow-up (the principal time point for evaluating child functioning), there was a small positive effect on expressive language (Effect size (ES) = 0.34 (95% CI (-0.12, 0.80)) $p = .13$). Although there was no evidence of a

benefit of the intervention on the other child outcomes, differences between intervention and control groups increased from the post intervention to follow-up assessment for all four child outcomes assessed at both time points (expressive language, attention, parent reported behavior difficulties and emotion regulation). Of these, the most notable was the effect size for expressive language, which increased by .38, from $ES = -0.04$ (95% CI (-0.40, 0.32)) at post, to $ES = 0.34$ (95% CI (-0.12, 0.80)) at follow-up.

There was no consistent evidence of moderating effects of other variables on the primary outcomes.

With regard to parenting during book-sharing, at the post intervention assessment, compared to control group parents, those in the intervention group were rated as significantly better on all measures, with effect sizes ranging from 0.35 to 1.09 (see Table 3a). Thus, intervention group parents were more sensitive to the child's interest and signals, provided more scaffolding for child language and cognition, and engaged in more talk about mental states; and there was more parent-child reciprocity during book-sharing. The largest effects were observed on the sensitivity and cognitive scaffolding measures ($ES = 1.09$ (95% CI (0.66, 1.52)), and $ES = 1.04$ (95% CI (0.72, 1.36)), respectively). For all these parenting variables, the improvements in intervention group parents were substantial (see Table S8 for Baseline and Post scores).

Despite these large effects on parenting, this ITT analysis showed no evidence of their indirect benefit to children's development.

Per protocol Intervention Group vs. Control Analysis

Effect sizes for comparisons between per protocol intervention and control populations are shown in Table 3b. Adjusted means for Primary outcomes are shown in Table 4, and those for secondary outcomes in Table S9. (Analyses of individual components of composite measures

are shown in Table S5). The exclusion of the 16 parents who failed to attend the requisite number of sessions led to effect sizes being larger than for the ITT comparison. As can be seen in Table 3b, a small, significant positive effect was observed at follow-up for expressive language (ES = 0.46 (95% CI (0.02, 0.91)), $p = 0.04$), and small, albeit non-significant, effects were found for language comprehension (ES = 0.22 (95% CI (-0.22, 0.66)), $p = .27$), and attention (ES = 0.34 (95% CI (-0.02, 0.70)), $p = .06$).

As for the comparisons between the ITT Intervention and control groups, larger differences between per protocol intervention and control groups were seen at follow-up compared to post intervention for all four measures that were assessed on both occasions, the most notable increase in effect size again being for expressive language and attention (i.e., an increase of .48 for expressive language (from ES = -0.02 (95% CI (-0.37, 0.33)) at post, to ES = 0.46 (95% CI (0.02, 0.91)), $p = 0.04$) at follow-up), and .33 for attention (from ES = 0.01 (95% CI (-0.38, 0.40)) at post, to ES = 0.34 (95% (CI (-0.02, 0.70)) at follow-up).

Results for parenting on all variables during book-sharing showed similar statistically significant positive effects as obtained for the ITT analysis, with the largest effects again observed on the sensitivity and cognitive scaffolding measures (see Table 3b).

Exploratory Analyses

Engaged Intervention vs. Control group analyses

Given that the comparisons between the engaged intervention group and controls were conducted as post hoc analyses, we considered only the four primary child outcomes (i.e., expressive language, comprehension, attention, and executive function). Adjusted means and effect sizes are shown in Table 4. It can be seen that effect sizes for the comparisons between the engaged intervention group and controls were larger than those for both ITT and per-protocol

populations. There was a medium effect for expressive language (ES = 0.56 (95% CI (0.07, 1.05)), $p = 0.03$), with smaller effects for language comprehension (ES = 0.38 (95% CI (0.01, 0.75)), $p = 0.04$), and attention (ES = 0.37 (95% CI (-0.06, 0.80)), $p = 0.09$). Executive function showed no benefit of intervention. As for the ITT and per protocol group comparisons, the difference between the engaged population and controls increased between post intervention and follow up for the two primary outcomes assessed on both occasions (see Table 4).

Discussion

Our dialogic book-sharing intervention aimed to benefit child development by improving parents' book-sharing. The great majority of participants attended the requisite number of sessions for the program and, indeed, of these per protocol participants, almost all attended all seven sessions (Mean = 6.8, SD .48). The intervention was well-received by the great majority of participants, they reported using the program techniques at home, and, during directly observed interactions, all dimensions of their book-sharing showed substantial benefit. Indeed, both parental sensitivity and cognitive scaffolding showed large intervention effects.

Regarding child development, the magnitude of benefit to language was, in the main, consistent with effects found in previous trials, albeit not reaching conventional levels of significance in all comparisons. Since this is likely to have been a function of inadequate statistical power, in our discussion below we focus on effects sizes rather than statistical significance (as recommended by Sullivan & Feinn, 2012). Thus, for expressive language, we found effects (Cohen's d) of 0.34 and 0.46, respectively, for ITT and per protocol analyses, compared to the average of 0.41 in the meta-analysis of Dowdall et al. (2020), and 0.21 in the recent study of Burgoyne et al. (2018). For receptive language, the effect size for our intervention was 0.22 in the per protocol population, compared to the 0.26 average in the

Dowdall meta-analysis. For those who were well-engaged, the impact of the intervention was increased to 0.56 for expressive language and 0.38 for receptive language. Regarding child attention, while there was no benefit of the intervention in the ITT population, the intervention effect size was 0.34 in the per-protocol analysis, rising to 0.37 in the engaged group. This is consistent with a recent South African DBS trial (Dowdall et al., 2021), which found an effect size for child attention of 0.39. These effect sizes are all in the range which, according to rules of thumb, are considered to be educationally significant (i.e., $d = .25$; Promising Practices Network, 2007; WhatWorks Clearing House, 2007). In contrast to these results for child language and attention, there was no evidence of a benefit of the DBS intervention to child executive function, social development, or emotional-behavior difficulties.

Several aspects of our study require comment. The first concerns our findings for child language where, despite being in line with those for other DBS trials, effect sizes were somewhat lower than we predicted on the basis of interventions structured similarly to our own (e.g., Vally et al., 2015). One possible explanation is that children may be particularly likely to benefit from DBS at an age when their language is just starting to take off and is on a rapid developmental trajectory. The fact that the children in our study were between six and eighteen months older than those in DBS intervention studies that achieved the largest effect sizes for expressive language (e.g., Huebner, 2000; Vally et al., 2015; Whitehurst et al., 1988), and beyond the period of steepest gains in language acquisition, might have attenuated intervention effects. This possibility is supported by longitudinal naturalistic studies showing the benefit to child language at age three-five years of parental reading at 14-24 months, measured as part of the home learning environment, is greater than that of concurrent reading (Raikes et al., 2006; Rodriguez & Tamis-LeMonda, 2011). Finally, it may be of significance that our findings indicated that

differences between our intervention and the control group increased over the period between the immediate post intervention assessment and the subsequent one, three to four months later, especially for language. This is consistent with the idea that, beyond the first two years, intervention effects may take longer to manifest, as the pace of child language development slows.

A second possible explanation for the effects of our intervention on child language being somewhat smaller than predicted is that all the families in our study were already receiving support through Children's Centers and, together with the Bookstart initiative, this might have reduced the potential for our intervention to make a more substantial impact. Without delivering our intervention to parents who did *not* receive Children's Center input or Bookstart support, it is not possible to formally evaluate this hypothesis. Nevertheless, it is important to note that public provision for parenting and child development in the UK provides a very different context from that pertaining in the earlier South African DBS trial, where particularly large gains in child language and attention were effected (Murray et al., 2016; Vally et al., 2015). This latter work was conducted in far more extreme conditions of deprivation than those obtaining in our UK study and, where, furthermore, Early Child Development Center support was either lacking or of poor quality, children's books were largely unavailable, and there was no background culture to support parent book-sharing. In this context, where baseline child functioning was extremely low, the potential for effecting large changes in child language was therefore considerable.

With regard to other areas of child development, the performance of our study sample at baseline was somewhat poorer on measures of behavior difficulties and prosocial behavior than that for general population samples in the UK

(<https://www.sdqinfo.org/norms/UK3yearNorm.html>) (the UK population mean score for SDQ

total difficulties is 7.3 (5.4), whereas for our intervention and control groups at baseline it was 11.49 (4.89) and 11.81(4.83), respectively; similarly, the UK mean score for SDQ Pro-social behavior is 8.1(1.8), whereas for our intervention and control groups it was 6.84 (1.86) and 7.26 (1.61), respectively). These relatively raised levels of difficulties are consistent with data for children in Reading in general, which has rather poorer performance than the national average on measures of socio-emotional functioning (<https://www.gov.uk/government/statistics/early-years-foundation-stage-profile-results-2018-to-2019>), and, in particular, with the fact that our study population was drawn from areas within Reading that were more deprived (see above SDQ website for deprivation effects on behavior difficulties). They are also consistent with research showing that the impact of Children’s Centers on these child outcomes is limited, and does not overcome the substantial effects of background risk (Sammons et al., 2015). Our dialogic book-sharing intervention, which provided support for families additional to routine Children’s Center input, included elements designed to address these areas of child development. Thus, it introduced ‘cognitive scaffolding’ techniques relevant to executive function skills, it encouraged talk about mental states and perspectives to promote social understanding, and it supported discussion of difficult family interactions of relevance to conduct difficulties. The question arises, therefore, why the intervention showed no significant benefit to these child outcomes. One possible explanation is that our program was not long enough to effect change in these areas of development. This seems unlikely, however, given that other, successful, interventions for these child outcomes have been of the same length, or even briefer than our own, including interventions with the sharing of picture books as a core element (Chacko et al., 2018; Howard et al., 2016; Lecce et al., 2014; Ornarghi et al., 2011). A more plausible explanation is suggested by differences in the content of our own intervention compared to that of these more successful

programs, where it is notable that explicit messaging, as well as specific focused exercises and practice in relevant skills, have been integrated into the process of book-sharing. For example, with regard to executive function, in the study of Howard et al. (2016), specific, repeated activities involving working memory, inhibition and shifting were embedded within the picture book story as obstacles that the child had to help the main character solve, under the intervenor's instruction. Similarly, the dialogic book-sharing intervention study of Chacko et al. (2018), where child behavior problems were additionally targeted, included not only the positive parenting principles common to our own intervention, but also explicit learning points for parents concerning negative parenting practices, illustrated by video examples of negative book-sharing for group discussion. Finally, regarding child social development, two particularly successful interventions for theory of mind abilities conducted by Lecce et al. (2014) and by Ornarghi et al. (2011), used structured questioning or games concerning the mental states of the storybook characters, including their false beliefs, with corrective feedback and explanations to further stimulate children's social understanding. Together, these diverse intervention studies suggest that, for dialogic book-sharing to support these wider dimensions of child development effectively, it may not be sufficient to facilitate general discussion of relevant themes, but likely requires augmentation with repeated, specifically focused interactions between adult and child that explicitly address the relevant processes.

An important aspect of our findings that requires comment is the role of parental engagement in the intervention. The majority of participants in our study engaged well and, as in previous studies of parenting interventions (e.g., Nix et al., 2009), such engagement was associated with especially good child outcome. Moreover, and again in line with previous research (e.g., Dumas et al., 2007; Gennetian et al., 2019; Nix et al., 2009; Waanders et al., 2007),

we found that poor engagement was associated with relatively higher levels of the socio-economic disadvantage that characterized the study population as a whole (i.e., those who did not engage well had a lower level of educational achievement, and lower income). An understanding of the mechanisms by which these socio-economic factors compromise engagement in parenting interventions is critical to modifying their delivery to improve child outcome. These mechanisms are complex (Gennetian et al., 2019), involving structural barriers (e.g., competing demands of work, absence of transport), attitudes and beliefs (e.g., concerning potential long-term benefits of parenting programs), and self-cognitions (e.g., concerning parenting efficacy). They are, moreover, likely to be culture-specific. For example, in work conducted in peri-urban settlement communities in South Africa, where extreme poverty and limited education are common (Cooper et al., 2014; Dowdall et al., 2021; Murray et al., 2016; Vally et al., 2015), there was a high level of parental motivation to engage with the book-sharing intervention. The reason for this appears to be that in South Africa, an intervention aimed at improving school readiness was especially valued because the parents considered school success to be a route out of inter-generational poverty. Such motivations appear harder to inspire in parents in the UK, who may, for example, consider that their children are already adequately provided for through the mainstream education system (Rabe, 2019), that they are too overburdened to share books with their child, and that what they do as parents will not stand to make a difference to their child's development.

A further possible factor that might have contributed to any limited engagement in the current efficacy study is that the intervention was delivered by research facilitators who were previously unacquainted with the parents, and who delivered the program in isolation from other support the parents received from the Children's Centers. There is good evidence that the quality

of relationship between parents and intervenors is important in influencing attendance and engagement (e.g., Robbins et al., 2003). Indeed, this element has been considered fundamental to the relatively greater success of Reach out and Read, compared to other book-gifting programs that do not provide the same continuity and regularity of contact with the same, already familiar professional (de Bondt et al., 2020). In sum, to improve engagement and achieve more effective delivery of the program, adjustments may be required regarding structural barriers to participation, and full account taken of parental perceptions concerning their own role as potential educators, the significance they accord to parenting interventions, and their relationship with program facilitators.

Strengths and limitations.

Our study had a number of strengths, including its RCT design, high participation and follow-up rates, comprehensive measures of child development, directly observed parenting, and strong analytic techniques. Limitations include the relatively small number of Children's Centers, with some associated baseline variation between groups, and this in turn necessitating inclusion in analyses of a covariate that had not been pre-specified, namely, the amount of book-sharing parents did at baseline. Finally, the fact that sampling was based on planned mid-range medium effect sizes left the study underpowered to show significance for some of the small to medium effects found.

Conclusion

Our seven-session dialogic book-sharing intervention, provided in addition to 'usual care' in UK Children's Centers, was associated with considerable benefit to parenting behavior during book-sharing, with large effects on sensitivity and cognitive scaffolding. Where parents attended the requisite number of training sessions, child expressive language also showed a positive impact of

the intervention, and this was especially so where parents were actively engaged in the training program. Similar effects, albeit somewhat weaker, obtained for child language comprehension and attention. Our results for child language add to an accumulating body of evidence from outside the UK for the benefits of dialogic book-sharing to this child outcome. Those for child attention are more novel, but are consistent with the previous work in South Africa. These findings indicate the potential value of adding a dialogic book-sharing program to routine UK Children's Center provision to enhance these aspects of parenting and child functioning in families experiencing socio-economic disadvantage. This could help narrow achievement gaps between children that are associated with socio-economic inequity (Marmot et al., 2020). Despite these positive effects on parenting and child development, our program had no impact on child executive function, social understanding, and behavior. We conclude, both from this absence of an intervention effect and from the findings of other trials, that the standard dialogic book-sharing program requires augmentation with specific intervention strategies to benefit these wider dimensions of child development.

References

- Abidin, R. R. (1995). *Parenting Stress Index Short Form* (3rd ed.). Odessa, FL: Psychological Assessment Resources
- Adrian, J., Clemente, R., Villanueva, L., Rieffe, C. (2005) Parent-child picture book reading, mothers' mental state language and children's theory of mind *Journal of Child Language*. 32, 673-686

- Adrian, J. E., Clemente, R. A., & Villanueva, L. (2007). Mothers' use of cognitive state verbs in picture-book reading and the development of children's understanding of mind: A longitudinal study. *Child Development, 78*(4), 1052-1067.
- Allen G. (2011). *Early Intervention: The Next Steps*, London: Cabinet Office, p23.
- Alloway T.P. (2007). Automated working memory assessment: Manual: Pearson.
- Altman DG. (1985) Comparability of randomised groups. *Statistician, 34*,125-36.
- ASPE (2014) *Research Brief 4: The Early Achievement and Development Gap*. Office of the Assistant Secretary for Planning and Evaluation; Office of Human Services Policy- U.S. Department of Health and Human Services
- Baptista, J., Osório, A., Martins, E.C., Castiajo, P., Barreto, A.L., Mateus, V., Soares, I., Martins, C. (2017) Maternal and Paternal Mental-state Talk and Executive Function in Preschool Children *Social Development, 26* (1), 129-145.
- Baron-Cohen, S. (1992). Out of sight or out of mind: another look at deception in autism. *Journal of Child Psychology and Psychiatry, 33*, 1141-55
- Benenson J.F., Pascoe, J., Radmore, N. (2007). Children's altruistic behaviour in the dictator game. *Evolution and Human Behaviour, 28*, 168-175.
- Boeg Thomsen, D., Theakston, A., Kandemirci, B., & Brandt, S. (2021). Do complement clauses really support false-belief reasoning? A longitudinal study with English-speaking 2- to 3-year-olds. *Developmental psychology, 57*(8), 1210–1227. doi 10.1037/dev0001012
- Brandt, S., Buttelmann, D., Lieven, E., & Tomasello, M. (2016). Children's understanding of first- and third-person perspectives in complement clauses and false-belief tasks. *Journal of experimental child psychology, 151*, 131–143. doi10.1016/j.jecp.2016.03.004

- Burgoyne, K., Gardner, R., Whiteley, H., Snowling, M.J., Hulme, C. (2018). Evaluation of a parent-delivered early language enrichment programme for pre-school children. *Journal of Child Psychology and Psychiatry*, 59(5), 545-555. doi: 10.1111/jcpp.12819.
- Bus, A. G., & van IJzendoorn, M. H. (1995). Mothers reading to their three-year-olds: The role of mother-child attachment security in becoming literate. *Reading Research Quarterly*, 30, 998–1015.
- Campbell, M. K., Mollison, J., Steen, N., Grimshaw, J. M., Eccles, M. (2000) Analysis of cluster randomized trials in primary care: a practical approach, *Family Practice*, 17(2), 192–196, <https://doi.org/10.1093/fampra/17.2.192>
- Centre for Social Justice (2014) *Closing The Divide: Tackling educational inequality in England*. September, 2014.
- Chacko, A., Fabiano, G. A., Doctoroff, G. L., & 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 Psychology* 47(1), 79–93. <https://doi.org/10.1080/15374416.2016.1266648J>
- Chen H., Lu Y., Slye N. (2020) Testing for baseline differences in clinical trials. *International Journal Clinical Trials* 7(2),150-3
- Cicchetti D, Rogosch FA, Toth SL. (1998) Maternal depressive disorder and contextual risk: contributions to the development of attachment insecurity and behavior problems in toddlerhood. *Developmental Psychopathology*, 10, 283-300
- Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic

- Cooper, P. J., Vally, Z., Cooper, H., Radford, T., Sharples, A., Tomlinson, M., & Murray, L. (2014). Promoting mother–infant book sharing and infant attention and language development in an impoverished South African population: a pilot study. *Early Childhood Education Journal, 42*, 143-152.
- Crain-Thoreson, C., Dahlin, M. P., & Powell, T. A. (2001). Parent–child interaction in three conversational contexts: variations in style and strategy. *New Directions for Child and Adolescent Development, 92*, 23-38.
- de Bondt, M., Willenberg, I. A., & Bus, A. G. (2020). Do book giveaway programs promote the home literacy environment and children’s literacy-related behavior and skills? *Review of Educational Research, 90*(3),349–375 <https://doi.org/10.3102/0034654320922140>
- Decety, J. (2015). The neural pathways, development and functions of empathy. *Current Opinion in Behavioral Sciences, 3*, 1-6
- Denham, S. A. (1986) Social Cognition, Prosocial Behavior, and Emotion in Preschoolers: Contextual Validation. *Child Development, 57*, 194-201.
- Department for Education (2011) *Evaluation of the Graduate Leader Fund Final report*. London: Department for Education.
- Department for Education (2014). *Early Years Pupil Premium and funding for two-year-olds*. London: Department for Education.
- <https://www.ncbi.nlm.nih.gov/pubmed/27883189>
- Devine, R.T., Hughes, C. (2018) Family correlates of False Belief Understanding in early childhood: a meta-analysis. *Child Development, 89*(3), 971-987. doi: 10.1111/cdev.12682

- Devine, R.T., Hughes, C. (2019) Let's Talk: Parents' Mental Talk (Not Mind-Mindedness or Mindreading Capacity) Predicts Children's False Belief Understanding *Child Development, 90*(4), 1236-1253
- Dowdall, N. (2015). The effectiveness of shared picture book reading interventions on child language outcomes: A systematic review and meta-analysis. Masters Thesis, Oxford University.
- Dowdall, N, Melendez-Torres G.J., Murray, L., Gardner, F., Hartford, L., Cooper, P. J. (2020). Shared Picture Book Reading Interventions for Child Language Development: A Systematic Review and Meta-Analysis *Child Development, 91*(2) 383-399
- Dowdall, N., Murray, L., Skeen, S., Marlow M., De Pascalis, L., Gardner, F., Tomlinson, M., Cooper, P. J. (2021) Book-sharing for parenting and child development in South Africa: a randomized controlled trial *Child Development, 00* 1-16 DOI: 10.1111/cdev.13619
- Dumas, J. E., Nissley-Tsiopinis, J., Moreland, A. D. (2007). From intent to enrollment, attendance, and participation in preventive parenting groups. *Journal of Child and Family Studies, 16*, 1–26
- Eldridge, S., & Kerry, S. (2012). *A Practical Guide to Cluster Randomised Trials in Health Services Research*. John Wiley & Sons.
- Ensor, R., Devine, R.T., Marks, A., Hughes, C. (2014) Mothers' Cognitive References to 2-Year-Olds Predict Theory of Mind at Ages 6 and 10 *Child Development, 85*(3) 1222-
- Fay-Stammach, T., Hawes, D. J., & Meredith, P. (2014). Parenting influences on executive function in early childhood: A review. *Child Development Perspectives, 8*, 258-264.

- Fenson, L., Pethick, S., Renda, C., Cox, J.L. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, 21, 95–116
- Fernald A, Weisleder A. (2011) Early language experience is vital to developing fluency in understanding. In: Neuman SB, Dickinson DK, editors. *Handbook of Early Literacy Research*. Vol. 3. New York: The Guilford Press; 2011. pp. 3–19
- Fletcher, K.L., Reese, E. (2005). Picture book reading with young children: A conceptual framework. *Developmental Review*, 25(1) 64-103.
- Freemantle, N., Calvert, M., Wood, J., Eastaugh, J., & Griffin, C. (2003). Composite outcomes in randomized trials: Greater precision but with greater uncertainty? *JAMA*, 289(19), 2554-2559.
- Garvey, C., Julion, W., Fogg, L., Kratovil, A., Gross, D. (2006). Measuring participation in a prevention trial with parents of young children. *Research in Nursing and Health*, 29, 212–222
- Gennetian, L. A., Marti, M., Lorenzo Kennedy, J., Jin Han K., Duch, H. (2019) Supporting parent engagement in a school readiness program: Experimental evidence applying insights from behavioral economics *Journal of Applied Developmental Psychology*, 62, 1-10, <https://doi.org/10.1016/j.appdev.2018.12.006>
- Goldman, D. Z., Shapiro, E. G., Nelson, C. A. (2004). Measurement of vigilance in 2-year-old children, *Developmental Neuropsychology*, 25(3), 227-250.
- Goldsmith, H., Rothbart, M. (1993) *The laboratory temperament assessment battery (LAB-TAB)*. University of Wisconsin.
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40, 1337-1345

- Gottfried, A.W., Schlackman, J., Eskeles Gottfried, A., Boutin-Martinez, A.S. (2015) Parental Provision of Early Literacy Environment as Related to Reading and Educational Outcomes Across the Academic Lifespan, *Parenting, 15*, 24-38
- Hall J., Sylva, K., Sammons, C., Smees, R., Evangelou, M., Smith, T., Goff, J. (2021) Investigating the reliability and validity of the Toddler Home Learning Environment (THLE) scale. doi: 10.3389/feduc.2021.581005
- Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11*, 49–76.
- Hoff, E. (2013) Interpreting the early language trajectories of children from low SES and language minority homes: implications for closing achievement gaps. *Developmental Psychology, 49*, 4–14. doi: 10.1037/a0027238.
- Hoff-Ginsberg, E. (1991). Mother–child conversations in different social classes and communicative settings. *Child Development, 62*, 782-796.
- Hofmann, S., G., Doan, S. N., Sprung, M., Wilson, A., Ebesutani, C., Andrews, L. A., Curtiss, J., Harris, P.L. (2016) Training children’s theory-of-mind: A meta-analysis of controlled studies, *Cognition, 150* 200-212 doi.10.1016/j.cognition.2016.01.006.
- Howard S, Melhuish, E. (2017). An Early Years Toolbox (EYT) for Assessing Early Executive Function, Language, Self-Regulation, and Social Development: Validity, Reliability, and Preliminary Norms. *Journal of Psychoeducational Assessment, 35*, 255-275
- Howard, S., Powell, T., Vasseleu, E., Johnstone, S., Melhuish, E. (2016) Enhancing preschoolers’ executive functions through embedding cognitive activities in shared book reading *Educational Psychology Review 28*(1) DOI 10.1007/s10648-016-9364-4

- Huebner, C. E., (2000) Promoting toddlers' language development through community-based intervention *Journal of Applied Developmental Psychology*, 21(5), 513–535
- Hughes, C., Daly, I., Foley, S., White, N., & Devine, R. T. (2015). Measuring the foundations of school readiness: Introducing a new questionnaire for teachers - The Brief Early Skills and Support Index (BESSI). *The British Journal of Educational Psychology*, 85(3), 332–356. <https://doi.org/10.1111/bjep.12076>
- Hughes, C., & Devine, R. T. (2019). For better or for worse? Positive and negative parental influences on young children's executive function. *Child Development*, 90(2), 593-609.
- Hughes, C., Devine., RT, Mesman, J., Blair, C. (2020) Understanding the terrible twos: A longitudinal investigation of the impact of early executive function and parent–child interactions. *Developmental Science*, 23:e12979. [doi.10.1111/desc.12979](https://doi.org/10.1111/desc.12979)
- Hughes, C., Ensor, R. (2007) Positive and protective: Effects of early theory of mind on problem behaviors in at-risk preschoolers. *Journal of Child Psychology and Psychiatry*, 48(10), 1025-1032.
- Hughes, C., Ensor, R., Marks, A. (2011) Individual differences in false belief understanding are stable from 3 to 6 years of age and predict children's mental state talk with school friends. *Journal of Experimental Child Psychology*, 108, 96-112.
- ICH Harmonised Tripartite Guideline. (1999). Statistical principles for clinical trials. *Statistics in Medicine*, 18(15), 1905-1942.
- Jaroslawska, A.J., Gathercole, S.E., Logie, M.R, Holmes, J. (2014). Following instructions in a virtual school: Does working memory play a role? *Memory & Cognition*, 44, 580-589.

- Justice, L.M., Jiang, H., Bates, R., Koury, A. (2020) Language disparities related to maternal education emerge by two years in a low-income sample *Maternal and Child Health Journal*, 24 (11), 1419-1427.
- Kannass, K. N., Oakes, L. M., & Shaddy, D. J. (2006). A longitudinal investigation of the development of attention and distractibility. *Journal of Cognition and Development.*, 7, 381-409.
- Khan, K.S., Purtell, K.M., Logan, J., et al. (2017) Association between television viewing and parent-child reading in the early home environment. *Journal of Developmental Behavioral Pediatrics*, 7(38) 521–527.
- Kochanska, G., Aksan, N. (1995). Mother-child mutually positive affect, the quality of child compliance to requests and prohibitions, and maternal control as correlates of early internalization. *Child Development*, 66(1):236-254.
- Landis, R., S., Beal, D. J., Tesluk, P.E. (2000) A comparison of approaches to forming composite measures in structural equation models *Organizational Research Methods*, 3(2), 186–207.
- Lecce, S., Bianco, F., Demicheli, P., Cavallini, E. (2014) Training preschoolers on first-order false belief understanding: transfer on advanced tom skills and metamemory. *Child Development*, 85(6) 2404–2418
- Leech, K. A., McNally, S., Daly, M., Corriveau, K.H. (2022) Unique effects of book-reading at 9-months on vocabulary development at 36-months: Insights from a nationally representative sample of Irish families. *Early Childhood Research Quarterly*, 58, 242-253,

- Lengua, L. J., Kiff, C., Moran, L., Zalewski, M., Thompson, S., Cortes, R., & Ruberry, E. (2014). Parenting mediates the effects of income and cumulative risk on the development of executive control. *Social Development, 23*, 631-649
- Logan, J.A.R., Justice, L.M., Yumuş, M., Chaparro-Moreno, L.J. (2019) When Children Are Not Read to at Home: The Million Word Gap *Journal of Developmental and Behavioral Pediatrics, 40* (5), 383-386.
- Marmot, M. Allen, J., Boyce, T., Goldblatt, P., Morrison, J. (2020) Health equity in England: the Marmot Review 10 years on <https://www.health.org.uk/publications/reports/the-marmot-review-10-years-on>.
- McGinty, A. S., Justice, L. M., Zucker, T. A., Gosse, C., & Skibbe, L. E. (2012). Shared-reading dynamics: Mothers' question use and the verbal participation of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 55*(4), 1039–1052. [https://doi.org/10.1044/1092-4388\(2011/10-0298\)](https://doi.org/10.1044/1092-4388(2011/10-0298))
- Melhuish, E.C., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., & Phan, M. (2008) Effects of the Home Learning Environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues, 64*, 95-114.
- Mingebach, T., Kamp-Becker, I., Christiansen, H., & Weber, L. (2018). Meta-meta-analysis on the effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. *PloS One, 13*(9), e0202855.
- Mol, S.E., Bus, A.G., de Jong, M.T., & Smeets, D.J. (2008). Added value of dialogic parent–child book readings: A meta-analysis. *Early Education and Development, 19*, 7-26.

- Morgan, P.L., Farkas, G., Hillemeier, M.M., Hammer, C.S. and Maczuga, S. (2015), 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, 86, 1351-370. doi.org/10.1111/cdev.12398
- Murray, L., De Pascalis, L., Tomlinson, M., Vally, Z., Dadomo, H., MacLachlan, B., Woodward, C., Cooper, P.J. (2016) Randomized controlled trial of a book-sharing intervention in a deprived South African community: effects on carer-infant interactions, and their relation to infant cognitive and socio-emotional outcome. *Journal of Child Psychology and Psychiatry*, 57, 1370-1379
- Murray, L., Halligan, S.L., Cooper, P. J. Postnatal Depression and Child Development (2019) In C. Zeanah (Ed.) *Handbook of Infant Mental Health 4th Edition* NY Guilford Press 172-186
- Murray, L., Jennings, S., Mortimer, A., Prout, A., Melhuish, E., Hughes, C., . . . Cooper, P. J. (2018). The impact of early-years provision in Children's Centres (EPICC) on child cognitive and socio-emotional development: Study protocol for a randomised control trial. *Trials*, 19, 450.
- Murray, L., Rayson, H., Ferrari, P-F, Wass, S. V., Cooper, P. J. (2022) Dialogic Book-Sharing as a Privileged Intersubjective Space *Frontiers in Psychology*, 13
DOI:10.3389/fpsyg.2022.786991
- Neitzel, C., & Stright, A. D. (2003). Mothers' scaffolding of children's problem solving: establishing a foundation of academic self-regulatory competence. *Journal of Family Psychology*, 17(1), 147-159.

- NICHD Early Child Care Research Network (1998). Early child care and self-control, compliance, and problem behavior at twenty-four and thirty-six months. *Child Development, 69*, 1145-1170
- Ninio, A. (1983). Joint book reading as a multiple vocabulary acquisition device. *Developmental Psychology, 19*, 445–451.
- Nix, R. L., Bierman, K. L., McMahon, R. J. (2009) How does attendance quality of participation affect treatment response to parent management training? *Journal of Consulting and Clinical Psychology, 77*(3), 429-438
- Noble, C., Cameron-Faulkner, T., & Lieven, E. (2018). Keeping it simple: The grammatical properties of shared book reading. *Journal of Child Language, 45*(3), 753-766.
doi:10.1017/S0305000917000447
- Obradovic, J., Yousafzai, A. K., Finch, J. E., & Rasheed, M. A. (2016). Maternal scaffolding and home stimulation: Key mediators of early intervention effects on children's cognitive development. *Developmental Psychology, 52*(9), 1409-1421
- Ornaghi, V., Brockmeier, J., & Grazzani, I. (2011). The role of language games in children's understanding of mental states: A training study. *Journal of Cognition and Development, 12*, 239–259. doi:10.1080/15248372.2011. 563487
- Pauli-Pott, U., Mann, C., Becker, K. (2021) Do cognitive interventions for preschoolers improve executive functions and reduce ADHD and externalizing symptoms? A meta-analysis of randomized controlled trials. *European Child and Adolescent Psychiatry, 30*(10), 1503-1521.
- Perry, R. E., Braren, S. H., Blair, C. (2018) Socioeconomic risk and school readiness: longitudinal mediation through children's social competence and executive function

- Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2018.01544>
- Peskin, J, Astington, J. (2004). The effects of adding metacognitive language to story texts. *Cognitive Development* 19, 253-273
- Petersen, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A., Pettit, G. S., & Van Hulle, C. A. (2013). Language ability predicts the development of behavior problems in children. *Journal of Abnormal Psychology*, 122(2), 542–557. <https://doi.org/10.1037/a0031963>
- Piotrowska, P.J., Tully, L.A., Lenroot, R. *et al.* (2017) Mothers, fathers, and parental systems: a conceptual model of parental engagement in programmes for child mental health—Connect, Attend, Participate, Enact (CAPE). *Clin Child Fam Psychol Rev* 20, 146–161 <https://doi.org/10.1007/s10567-016-0219-9>
- Pocock, S. J. Simon, R (1975). Sequential Treatment Assignment with Balancing for Prognostic Factors in the Controlled Clinical Trial. *Biometrics. International Biometric Society*. 31, 103–115.
- Promising Practices Network. (2007). How programs are considered. <http://www.promisingpractices.net>
- Rabe, B. (2019). Do school inputs crowd out parents' investments in their children? *IZA World of Labor* doi: 10.15185/izawol 460
- Raikes, H., Pan, B. A., Luze, G., Tamis-LeMonde, C. S., Brooks-Gunn, J., Banks Tarullo, J. C., et al. (2006). Mother-child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development*, 6, 193–235.
- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, 33, 12–21.

- Robbins, M. S., Turner, C. W., Alexander, J. F., Perez, G. A. (2003) Alliance and dropout in family therapy for adolescents with behavior problems: Individual and systemic effects. *Journal of Family Psychology, 17*, 534–544.
- Robinson, J.L., Emde, R. N. (2004) Mental health moderators of Early Head Start on parenting and child development: maternal depression and relationship attitudes. *Parenting: Science and Practice, 4*(1) 73–97
- Rodrigues A.M., O'Brien, N., French, D.P., Glidewell, L., Sniehotta, F.F. (2015). The question-behavior effect: genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. *Health Psychology, 34* 61–78.
- Rodriguez, E.T., Tamis-LeMonda, C. S. (2011) Trajectories of the home learning environment across the first 5 years: associations with children's vocabulary and literacy skills at prekindergarten *Child Development, 82* (4) 1058–1075
- Roheger, M., Hranovska, K., Martin, A. K., Meinz, M. (2022) A systematic review and meta-analysis of social cognition training success across the healthy lifespan *Scientific Reports 12*(1):3544 DOI:[10.1038/s41598-022-07420-z](https://doi.org/10.1038/s41598-022-07420-z)
- Roman, G.D., Ensor, R., Hughes, C. (2016) Does executive function mediate the path from mothers' depressive symptoms to young children's problem behaviors? *Journal of Experimental Child Psychology, 142*, 158-170.
- Ryan R.,M., Fauth, R.,C., Brooks-Gunn, J. (2006) Childhood poverty: Implications for school readiness and early childhood education. In: Spodek B., Saracho O.N., editors. *Handbook of research on the education of children*. 2nd ed Erlbaum; Mahwah, NJ: pp. 323–346.
- Salo, V-C. Rowe, M., Leech, K.A., Cabrera, N.J. (2016) Low-income fathers' speech to toddlers during book reading versus toy play. *Journal of Child Language, 43*, 1385-1399

- Sammons, P., Toth, K., Sylva, K., Melhuish, E., Siraj, I., & Taggart, B. L. (2015). The long-term role of the home learning environment in shaping students' academic attainment in secondary school. *Journal of Children's Services, 10*, 189-201.
- Schulz K.F., Grimes, D.A. (2005). Multiplicity in randomized trials I: endpoints and treatments. *Lancet, 365*, 1591–95.
- Scott N.W., McPherson, G.C., Ramsay, C.R., Campbell, M.K. (2002) The method of minimization for allocation to clinical trials. a review. *Control Clin Trials. 23(6):662-74*. doi: 10.1016/s0197-2456(02)00242-8. PMID: 12505244.
- Sénéchal, M., Thomas, E., & Monker, J. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87*, 218–229
- Sénéchal, M., Cornell, E.H., Broda, L.S. (1995) Age-related differences in the organization of parent-infant interactions during picture-book reading, *Early Childhood Research Quarterly, 10(3)* 317-337 [doi 10.1016/0885-2006\(95\)90010-1](https://doi.org/10.1016/0885-2006(95)90010-1).
- Senn S. Testing for baseline balance in clinical trials. (1994) *Stat Med. 13:1715–1726*
- Snow, C. E. & Ferguson, D. (eds) (1977) *Talking to children: language input and acquisition*. New York: C.U.P
- Song, M. K., Lin, F. C., Ward, S. E., & Fine, J. P. (2013). Composite variables: When and how. *Nursing Research, 62(1)*, 45–49.
- Sylva, K. (2014). The role of families and pre-school in educational disadvantage. *Oxford Review of Education, 40*, 680-695

- Sullivan, G., M., Feinn, R. (2012) Using effect size—or why the P value is not enough. *Journal of Graduate Medical Education*, 4(3): 279–282. <https://doi.org/10.4300/JGME-D-12-00156.1>
- Rodriguez, E.T., Tamis-LeMonda, C. S., Spellmann, M. E. et al., (2009) The formative role of home literacy experiences across the first three years of life in children from low-income families. *Journal of Applied Developmental Psychology*, 30 (6), 677–694
- Taumoepeau, M. (2016) Maternal Expansions of Child Language Relate to Growth in Children’s Vocabulary. *Language Learning and Development*, 12(4), 429-446
- Treasure T., MacRae K. D. (1998) Minimisation: the platinum standard for trials? *BMJ*, 317, 362–363.
- Vygotsky, L. (1978). *Mind in Society: the development of higher psychological processes* Cambridge, Massachusetts: Harvard University Press.
- Vally, Z., Murray, L., Tomlinson, M., Cooper, P.J. (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, 865-873.
- Waanders, C., Mendez, J. L., Downer, J. T. (2007) Parent characteristics, economic stress and neighborhood context as predictors of parent involvement in preschool children's education. *Journal of School Psychology*, 45(6) 619-636
doi.org/10.1016/j.jsp.2007.07.003
- Wechsler, D. (2012). Wechsler preschool and primary scale of intelligence – Fourth edition: Manual: Pearson.
- Weissman, M. M., Wickramaratne, P., Nomura, Y., Warner, M., Pilowsky, D., Verdelli, H. (2006) Offspring of depressed parents: 20 years later. *American Journal of Psychiatry*

163(6):1001-8 DOI:[10.1176/appi.ajp.163.6.1001](https://doi.org/10.1176/appi.ajp.163.6.1001)

Weissman, M., M. Pilowsky, D., et al. (2006) Remissions in Maternal Depression and Child

Psychopathology A STAR*D-Child Report *JAMA*. 295:1389-1398

Wellman, H., Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75 523-541.

What Works Clearing House. (2014). Procedures and Standards Handbook.

<https://ies.ed.gov/ncee/wwc/Handbooks#procedures>

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, 552–559.

Whitehurst, G. J., and Lonigan, C. J. (1998) Child Development and Emergent Literacy. *Child Development*, 69, 848–872.

Wiig, E.H, Secord, W.A., Semel, E. (2004). Clinical evaluation of language fundamentals—Preschool, second edition (CELF Preschool-2). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment

Wimmer, H., Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128

Zigmond, A.S., Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67, 361-70.

Table 1*The Book-sharing Intervention Structure and Content*

Session and Book used	Theme	Content
1. 'Handa's Surprise' E. Browne	Introduction and core skills	Importance of book-sharing; setting up routines; following child's lead, asking questions to engage child, linking content to child experience
2. 'Little Helpers' L. Murray & P. Cooper	Elaborating and linking	Picking up on child's focus of interest and elaborating on it. Making links between the book content and the child's wider experience.
3. 'Handa's Hen' E. Browne	Numbers and comparisons	Practicing activities that promote enumerating and making comparisons, working memory, inhibition and shifting.
4. 'Hug' J. Alborough	Emotions	Talking about the book characters' feelings. Naming and contextualizing them. Linking the book characters' feelings to child's experience
5. 'Harry the Dirty Dog' G. Zimmerman	Intentions	Discussing book characters' desires, intentions and beliefs and why they might behave as they do.
6. 'Harry by the Sea' G. Zimmerman	Perspectives	Highlighting that different book characters can have differing desires, intentions, beliefs and perspectives, and how these influence their behavior.
7. 'The wrong side of the bed' E. Ardizzone	Relationships, Summary and taking forward.	Discussing everyday family relationships, including conflicts and their resolution. Summary of key learning points from sessions 1-6; and discussion of continuing with regular book-sharing (including accessing books).

Booster Books sent out after the intervention sessions: 'No Roses for Harry' by G. Zimmerman; 'Not Now, Bernard' by D. McKee

Table 2.*Key Demographic Characteristics of Different Study Groups*

Demographic Characteristic	Control n = 108	ITT intervention n = 110	Per protocol intervention n = 94	Non-per protocol intervention n = 16	Engaged intervention n = 65	Not engaged intervention n = 45
Child age (months) M (SD)	35.3 (5.65)	33.2 (5.02)	33.2 (5.01)	33.4(5.20)	33.1 (4.95)	33.4 (5.17)
Maternal age (years) M (SD)	33.7 (6.70)	34.2 (5.65)	34.7 (5.43)	31.3(6.23)	34.8 (5.32)	33.3 (6.05)
Child male %	63.0	53.6	56.4	37.5	58.5	46.7
Parent Ethnicity %						
Asian	8.3	20.0	21.3	12.5	15.4	26.7
Black	6.5	13.6	12.8	18.8	12.3	15.6
White	83.3	63.6	62.8	68.8	70.8	53.3
Multilingual %	19.4	39.1	41.5	25.0	43.1	33.3
Parent Education %						
</= GCSE (US Grade 10)	17.6	16.4	10.6	50.0	7.7	28.9
</= A Level (US Grade 12) or equivalent post-school qualification	38.9	42.8	35.1	87.5	35.4	53.3
Income %						
< £16,000	13.0	22.7	18.1	50.0	15.4	33.3
< £25,000	27.8	37.2	31.9	68.8	32.3	44.4
Baseline book-sharing (minutes per week) M (SD)	n = 98 114.6(70.71)	n = 97 71.2(63.89)	n = 87 72.0(64.34)	n = 10 64.2(62.61)	n = 61 76.6(69.30)	n = 36 61.9(53.12)

Table 3.

Summary results for all study outcomes for a) Intervention (I) vs. Control (C) ITT; and b) Per protocol vs Control

Variables	Time point	a) Intervention vs Control (ITT)		b) Intervention vs Control (Per Protocol)	
		Standardized effect size (I-C) (95% CI) ^a	<i>P</i>	Standardized effect size (I-C) (95% CI) ^a	<i>P</i>
Primary Child Outcomes					
Expressive language (EYT)#	Post	-0.04(-0.40, 0.32)	0.81	-0.02 (-0.37, 0.33)	0.91
	Follow-up	0.34 (-0.12, 0.80)	0.13	0.46 (0.02, 0.91)	0.04
Language comprehension (CELF)*	Follow-up	0.07 (-0.28, 0.43)	0.69	0.22 (-0.22, 0.66)	0.27
Attention composite	Post	-0.02 (-0.39, 0.34)	0.89	0.01 (-0.38, 0.40)	0.97
	Follow-up	0.11 (-0.24, 0.47)	0.53	0.34 (-0.02, 0.70)	0.06
Executive function composite	Follow-up	-0.19 (-0.69, 0.31)	0.41	-0.01 (-0.53, 0.52)	0.98
Secondary Child Outcomes					
Social development composite	Follow-up	0.04 (-0.48, 0.56)	0.87	0.08 (-0.35, 0.52)	0.68
Behavior Difficulties					
Defiance (Don't touch)(low good)^	Follow-up	-0.17 (-0.47, 0.14)	0.28	-0.16 (-0.47, 0.16)	0.32
Total difficulties (SDQ) (low good)	Post	0.01 (-0.26, 0.27)	0.96	0.02 (-0.25, 0.29)	0.88
	Follow-up	-0.11 (-0.44, 0.23)	0.53	-0.17 (-0.52, 0.18)	0.34
Emotion Regulation					
CSBQ (low good)	Post	0.04 (-0.42, 0.49)	0.86	0.02 (-0.38, 0.41)	0.94
	Follow-up	-0.08 (-0.50, 0.33)	0.67	-0.13 (-0.47, 0.21)	0.43

Parent Book-Sharing					
Sensitivity	Post	1.09 (0.66, 1.52)	<0.001	1.21 (0.74, 1.67)	<0.001
Scale Reciprocity	Post	0.77 (0.39, 1.15)	0.001	0.79 (0.38, 1.19)	0.001
Event Reciprocity	Post	0.35 (0.03, 0.67)	0.030	0.35 (0.03, 0.68)	0.032
Cognitive scaffolding	Post	1.04 (0.72, 1.36)	<0.001	1.10 (0.77, 1.43)	<0.001
		Adjusted ratio (95% CI) I/C		Adjusted ratio (95% CI) I/C	
Mental state talk	Post	2.93 (2.13, 4.05)	<0.001	3.31 (2.26, 4.29)	<0.001

Note. Adjusted means for primary outcomes are in Table 4, and for secondary outcomes in S6 for ITT population and S9 for per protocol population. Multiple imputation was used for missing data for the primary outcomes and social development composite for the ITT analyses only. ^aAdjusting for baseline (and baseline x time point, where appropriate), child age, gender, multilingualism and parental education. For Executive Function and Social Development composites, models used available components at baseline. [#]adjusting for base-line book-sharing (S10 shows non-adjusted analyses). ^{*}adjusting for baseline EYT. [^]square root transformation. Intra-cluster correlation = 0 to 0.07 for all measures. Where a low value is good, a negative standardized ES reflects a positive effect of intervention.

Table 4.

Primary child outcomes adjusted mean scores for the ITT, Per protocol and Engaged groups, with intervention effects for the Engaged group

Variable	Time point	Intervention n	Control n	Intervention adjusted mean (95% CI)	Control adjusted mean (95% CI)	Adjusted difference (95% CI)
Intention to treat						
Expressive Language EYT#	Post	110	108	21.93 (20.73, 23.12)	22.12 (20.89, 23.35)	-0.20 (-1.95, 1.56)
	Follow - up	110	108	28.13 (26.79, 29.47)	26.72 (25.42, 28.01)	1.41 (-0.49, 3.31)
Comprehension CELF*	Follow-up	110	108	56.94 (54.21, 59.66)	56.15 (53.48, 58.81)	0.79 (-3.21, 4.79)
Attention composite	Post	110	108	0.20 (0.07, 0.34)	0.21 (0.10, 0.33)	-0.01 (-0.19, 0.16)
	Follow-up	110	108	0.41 (0.27, 0.54)	0.35 (0.22, 0.47)	0.06 (-0.12, 0.24)
Executive function composite	Follow-up	110	108	-0.13 (-0.38, 0.12)	0.00 (-0.24, 0.25)	-0.13 (-0.49, 0.22)
Per protocol						
Expressive Language EYT#	Post	86	96	22.76 (21.63, 23.89)	22.85 (21.71, 23.98)	-0.08 (-1.72, 1.56)
	Follow - up	85	98	29.25 (28.02, 30.48)	27.42 (26.21, 28.63)	1.83 (0.06, 3.59)
Comprehension CELF*	Follow-up	83	100	59.99 (56.58, 63.40)	57.56 (54.10, 61.02)	2.43 (-2.45, 7.31)
Attention composite	Post	58	79	0.22 (0.11, 0.33)	0.22 (0.12, 0.31)	0.00 (-0.14, 0.15)
	Follow-up	56	78	0.45 (0.35, 0.56)	0.32 (0.23, 0.41)	0.13 (-0.01, 0.28)
Executive function composite	Follow-up	71	90	0.06 (-0.17, 0.28)	0.06 (-0.16, 0.28)	-0.00 (-0.32, 0.31)
Engaged group						

							Standardized effect size (95% CI)	<i>P</i>
Expressive Language EYT#	Post	60	96	23.17 (21.81, 24.52)	23.07 (21.81, 24.32)	0.10 (-1.77, 1.97)	0.02 (-0.40, 0.44)	0.91
	Follow-up	60	98	29.79 (28.35, 31.22)	27.56 (26.28, 28.84)	2.22 (0.27, 4.17)	0.56 (0.07, 1.05)	0.03
Comprehension CELF*	Follow-up	58	100	61.79 (58.75, 64.83)	57.73 (55.47, 60.00)	4.06 (0.12, 8.00)	0.38 (0.01, 0.75)	0.04
Attention Composite	Post	34	79	0.23 (0.09, 0.36)	0.23 (0.13, 0.32)	-0.00 (-0.17, 0.17)	-0.01 (-0.44, 0.43)	0.98
	Follow-up	36	78	0.47 (0.33, 0.61)	0.33 (0.23, 0.42)	0.15 (-0.02, 0.32)	0.37 (-0.06, 0.80)	0.09
Executive Function Composite	Follow-up	50	90	0.10 (-0.18, 0.38)	0.08 (-0.18, 0.35)	0.01 (-0.37, 0.40)	0.02 (-0.59, 0.63)	0.94

Note. Multiple imputation used for the ITT population analysis of the primary outcomes. Adjusting for available baselines (and baseline by visit interaction where appropriate), child age, gender, multilingualism and parental education. #adjusting for base-line book-sharing. *adjusting for baseline EYT.

Figure 1

CONSORT

