The Early-years Provision In Children's Centres (EPICC) Study
June 2019

Lynne Murray, Susie Jennings, Hayley Perry, Xiaonan Liu, Claire Hughes, Edward Melhuish, Peter Cooper

The Research Team

Principal Investigators: Lynne Murray, Peter Cooper, Claire Hughes, Edward Melhuish
Trial Manager: Susie Jennings
Facilitators: Martin Andrews, Kath de Wilde
Statisticians: Hayley Perry, Xiaonan Liu
Research Assistants: Amber Prout, Alicia Mortimer, Emily Phillips, Chloe Rock, Hallam Cooper, Ola Lowicka
Collaborators: John Duncan, Joni Holmes
Children's Centres Team Manager Reading Borough Council: Corinne Dishington

Acknowledgements

We thank the parents and children who participated in the study, and the Children's Centres' staff who facilitated the intervention's implementation and provided facilities for child assessments. We thank student volunteers who helped with data coding and transcribing, and with child care. In particular, we thank Raveena Rai and Jade Orpwood for assistance with data management. We thank Jack Dainty for work on randomisation of the Children's Centres.

We thank the Nuffield Foundation for funding this project, and are particularly grateful to Ruth Maisey for her guidance and support.

The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics and the Ada Lovelace Institute. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org

List of Acronyms used:

CELF	Clinical Evaluation of Language Fundamentals
CI	Confidence Interval
CSBQ	Children's Social Behavior Questionnaire
DBS	Dialogic book-sharing
DfE	Department for Education
DVQ	Discipline and Violence Questionnaire
ECVT	Early Child Vigilance Task
EPICC	The Early-years Provision In Children's Centres Study
EYT	Early Years Tool-Box
HADS	Hospital Anxiety and Depression Scale
ICC	Intra-cluster correlations
IDACI	Income Deprivation Affecting Children Index
ITT	Intention to Treat
LAB-TAB	Laboratory Temperament Assessment Battery
LSOA	Lower Super Output Areas
PSI	Parental Stress Index
SDQ	Strengths and Difficulties Questionnaire

Executive summary

Background. A number of programmes have been developed to support parents' use of 'Dialogic Reading', or 'Dialogic Book-sharing' (DBS) with their children. The DBS method is based on extensive observational research on the kinds of parent-child interactions that best promote child development, and particularly language. It refers to a particular way of using a book with a child that sensitively follows and supports the child's interests, and engages them actively in a reciprocal interaction. Thus, it is not about an adult simply *reading* a book to a passive child who just listens.

Parents vary in the quality of their book-sharing interactions and, notably, those who are of lower SES and/or education are less likely to use beneficial dialogic techniques. Nevertheless, trials of DBS interventions, mainly conducted in the US, show that parents can be trained to use good dialogic book-sharing practices, with positive outcomes for their children. Little research has been conducted on DBS in the UK.

Aims. We conducted a cluster randomized controlled trial to determine whether, compared to normal Children's Centre input, a DBS intervention, conducted with parents of children aged 2-4 years in Children's Centres in Reading, UK, was associated with better child developmental outcome and parenting. The primary outcome was child cognitive development (measured by expressive language, comprehension, attention, and executive function); secondary outcomes were child social development, behaviour problems and emotion regulation, and parenting during book-sharing and during a child compliance task.

Methods. 110 participants were randomised to the DBS Intervention, and 108 to the Control group. The intervention was delivered in Children's Centres once a week for seven weeks by research facilitators. It was delivered to small groups of parents for 50 minutes, and after each group session, individual parents received support for five-ten minutes.

Assessments by researchers, unaware of whether or not the parents had received the intervention, were made of parenting and child development before the intervention started, shortly after the intervention period (Post assessment), and at Follow-up 4-6 months after the end of the intervention. 102 (93%) of Intervention group participants, and 105 (97%) of Control group participants were retained in the study.

Analyses. The main analyses were conducted on an Intention to Treat (ITT) basis; that is, results were considered for *all* participants who were recruited into the study, regardless of whether or not parents ultimately took part in the intervention. However, as we wanted to know whether the intervention was effective if it was actually delivered, we conducted secondary 'Per protocol' analyses (i.e., on the 94 of the 110 (85%) participants who were originally enrolled in the intervention group who attended intervention sessions). Finally, we were also concerned to know whether parents' positive engagement with the intervention (rather than simple attendance) influenced the results. We therefore conducted post hoc analyses for the primary child outcomes, using results from the 'engaged' participants (69% of those who attended).

Results. Both ITT and Per protocol analyses showed substantial and significant benefits of the intervention compared to control on book-sharing behaviour. 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. There were, however, no benefits of the intervention on parenting during the child compliance task.

We found no significant benefits of the intervention to any child outcome for the ITT analyses at either Post or Follow-up assessments, although there was a small, non-significant positive effect on expressive language at Follow-up (Effect size (ES) = 0.34 (95% CI (-0.12, 0.80)) $p = 0.13$). Results were generally better for the Per-protocol vs. the ITT population, with a small significant effect for expressive language at Follow-up (ES = 0.46 (95% CI (0.02, 0.91)), $p = 0.04$), and small, albeit non-significant, effects for language comprehension (ES = 0.22 (95% CI (-0.22, 0.66)), $p = 0.27$), and attention ES = 0.34 (95% CI (-0.02, 0.70)), $p = 0.06$). There was also an indication of greater benefit to children whose parents engaged with the intervention, with a medium effect for expressive language (ES = 0.56 (95% CI (0.07, 1.05)), $p = 0.03$), and 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)), $p = 0.09$).

For both ITT and Per-protocol populations, intervention effects increased between the Post intervention and Follow-up assessments on all four measures assessed on both occasions (expressive language, attention, parent reported behaviour difficulties and emotion regulation), the most notable being for expressive language and attention. Similarly, the

difference between the engaged population and controls increased between Post intervention and Follow-up for the two primary outcomes assessed on both occasions.

Interpretation. The intervention was highly successful in improving parental behaviour during book-sharing. Gains to child development were more limited, although, for language and attention, they were within the range found for other interventions using DBS, and of an order considered educationally significant (Promising Practices Network, 2007; WhatWorks Clearing House, 2007). There are a number of possible independent, and overlapping, explanations for there not being greater benefits to child outcome. First, the fact that the difference between Intervention and Control groups generally increased between Post and Follow-up assessments may mean that benefits to child development take more time to emerge.

Second, it is possible that the routine support provided to pre-school children in England (including Children's Centres' input and schemes providing books to families) is of sufficient quantity and quality that there was less scope for any additional benefits of our intervention to occur compared to some other contexts that lack the level of provision for child development that is available in the UK.

Third, with regard to child socio-emotional and behavioural outcomes, it is possible that standard DBS would need to be augmented with specific additional components for it to effect change in these areas.

Finally, despite significant benefits to parents' book-sharing for the Intervention group as a whole, not all parents perceived the intervention as being of value, and those who failed to engage may not have been motivated to implement it, with the consequence that their child could not benefit. Socio-economic and cultural factors appear to be relevant to such parental responses to the intervention. For example, compared to those who did engage well, those who did not tended to be less highly educated, and more often on very low incomes. The fact that the same type of intervention has been highly successful in settings where low education and poverty are common (e.g. in South Africa), suggests that other factors are also relevant, and that it is important to understand more about the acceptability of the intervention and its perceived relevance in different groups, For example, some parents may consider that their

children are already adequately provided for through the mainstream education system (Rabe, 2019), or feel themselves too overburdened to share books regularly with their child, or they may believe that what they do as parents may not stand to make a difference to their child's developmental progress.

Conclusion. A seven-session Dialogic book-sharing intervention was associated with substantial benefits to parenting behaviour during book-sharing. Gains in child development were more limited, comprising small-to-medium sized benefits to language and attention. Further investigation is required to establish whether further child benefits emerge in the longer term (and a follow up of the current sample is currently underway). Since the level of parental engagement in the intervention appeared to be important to child outcome, a greater understanding of parental perceptions may be required in order to deliver the intervention more effectively.

Background

There are marked disparities between pre-school children in key skills affecting school readiness (e.g., language, attention, managing behaviour and emotions, and social relationships) (DfE, 2014; Sutton Trust & Oxford University, 2014). Much of this disparity is linked to socio-economic disadvantage and its impact on the home learning environment (Sylva, 2014; Sammons et al., 2015; Melhuish et al., 2008). These early childhood effects of disadvantage are important, as they commonly persist and influence children's later academic achievements, employment and adjustment, thereby perpetuating inter-generational cycles of disadvantage (Centre for Social Justice, 2014; Allen, 2011). Children's Centres represent an ideal context in which to implement and evaluate programmes that could address this problem, as they provide for families from the antenatal period up to age five, and aim to promote parenting skills, as well as providing care for children. Research shows that such preschool provision can be of particular benefit to children's later performance and functioning at school when staff are highly trained and support parents' involvement in their children's learning at home (DfE, 2011). One parenting practice that stands to be of particular benefit to children's development is reading to them. For example, one longitudinal study that examined children's development from 3 years into adulthood found that the amount of maternal reading to the child in the preschool years was a significant predictor of their eventual educational attainment, over and above the influence of the mother's own education; the amount of reading *material* in the home, by contrast, was not associated with child and adult educational progress (Gottfried, Schlackman, Eskeles-Gottfried, Boutin-Martinez, 2015).[1] In fact, research has accumulated to show that a particular form of parental reading, 'Dialogic reading' or 'dialogic book-sharing', is particularly helpful to child development. This is a method of supporting an infant or young child with a picture book in a way that

[1] See also Law et al. (2018) for a recent review of parent-child reading interventions.

sensitively follows and supports the child's interests and engages them actively in a reciprocal interaction. Thus, it does not consist in simply reading to the child, but typically uses techniques to encourage the child's active participation. These include building on the child's interest and asking them questions in a developmentally appropriate manner (i.e., pitched just at their level of competence, or what Vygotsky termed their 'zone of proximal development') (Vygotsky, 1978), and linking the book content to the child's own experience. A review of 16 studies in 2008 (Mol, Bus, de Jong, & Smeets, 2008) showed that parental dialogic reading with children was of significantly greater benefit to children's language development (particularly expressive language) than 'normal reading'; and a more recent meta-analysis of 19 RCT's (Dowdall et al., 2019) confirmed this conclusion. Importantly, disadvantaged parents are less likely than others to share books with their children, and when they do, they tend not to use the 'dialogic' techniques that have been found particularly helpful to child development (Bus & van IJzendoorn, 1995; Heath, 1982; Raikes, Pan, Luze, Tamis-LeMonde, Brooks-Gunn, Banks Tarullo, 2006; Fletcher and Reese, 2005), and therefore they are especially in need of supportive guidance in using these techniques. In fact, we showed, in a highly disadvantaged South Africa population, that a six to eight-week book-sharing training programme brought about significant benefits to child attention and language (Cooper et al. 2014; Vally et al, 2015), benefits that were mediated by gains in parental sensitivity and reciprocity whilst sharing picture books (Murray et al., 2016). Importantly, despite the accumulation of 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, and one of our aims was to address this gap.[2]

[2] The first UK trial including Dialogic book-sharing, as just one component of a multi-component intervention, was published in 2018 (Burgoyne et al., 2018)

The EPICC study

Aims

The aims of the current UK study were to determine, via a cluster randomised controlled trial within Children's Centres, the impact of providing carers with training in supportive, dialogic, book-sharing with their young children. In particular, we aimed to determine the impact of the training on child cognitive development (primary outcome), social development, behaviour problems and emotion regulation (secondary outcomes), and on parenting (secondary outcome).

Context

The study was based in Reading Children's Centres. The population of Reading is representative of the general UK population, including its multi-ethnic communities and significant areas of deprivation. Over three thousand Reading children under 5 years live within the 20% Most Deprived Lower Super Output Areas (LSOAs) [National IDACI index], representing 25% of all Reading children of this age. Most (10 of 13) Children's Centres in Reading are situated in these deprived areas, and attendance from this group, recorded on the Reading Council data base, has consistently been over 70% of registered children for all Children's Centres. Inspection of the characteristics of the wards in which the Centres are based showed one of the 13 to be an outlier in terms of the Index of Multiple Deprivation (IMD) and ethnic group composition, and this Centre was excluded from the study. The remaining 12 Centres were randomly assigned either to the Intervention condition (6 centres) where parents received training in book-sharing, or to the Control condition (6 centres) where parents received normal Children's Centre input. (We did not randomize families with Centres to Intervention or Control conditions, as this posed a serious risk of contamination between Intervention and Control groups). Randomisation to the Intervention and Control

clusters was undertaken by an independent statistician, with minimisation on the IMD and the ethnic profile of the wards in which each of the Centres was located.

The Book-sharing Intervention

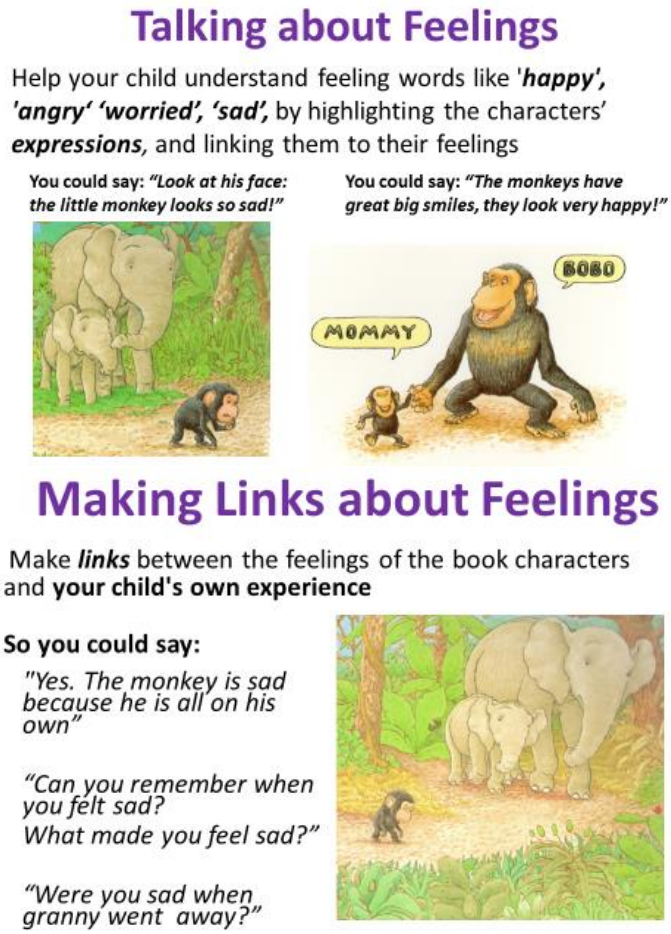
The intervention was an adaptation of one originally developed in the US by Whitehurst and colleagues (1988) and shown to be of benefit in numerous studies, principally conducted in the US, to child language development (Dowdall et al, 2019). The version we used was modelled on the one we had previously successfully delivered in South Africa (Cooper et al., 2014; Vally et al., 2015; Murray et al., 2016). 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. Flexible responsiveness is emphasised 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 Centres by two trained research facilitators who had a background of working in Children's Centres/the Early Years context. It comprised seven, weekly, facilitator-led small-group sessions with parents, lasting approximately 50 minutes, each one being followed by a brief period of one-to-one discussion between facilitator and each participating parent. During the group session, children were cared for in an adjoining play space by Children's Centre or study staff. The intervention provided parents with guidance on how to share books with their child in ways that best support child development. It was manualised, and delivered using a Power point format, including video examples of parents using good book-sharing practices with children. Each session focussed on a particular theme using a 'book of the week' to illustrate it, and covered specific book-sharing techniques to enhance child development. Books were selected to match each theme and to provide good opportunities to rehearse the sessions' techniques. They were either wordless, or text-light books, since this format has been shown to elicit more responsive, dialogic parent-child interactions and to afford more elaborated talk,

including talk concerning the mental states of book-characters (Sénéchal et al., 1995; Peskin and Astington, 2004; Noble et al., 2018). Each book of the week was given to participants at the end of the session, along with brief card containing reminders of the session's main points, and participants were encouraged to practise sharing the book with their child regularly over the coming week. At approximately two and three to four months after the Centre-based intervention, parents were sent an additional book as a booster, with accompanying notes on how it might be used with the child. The session themes and books are shown in Table 1, and Sample Power point slides from Session 4 are shown in Figure 1.

Table 1 The Book-sharing Intervention Content

***Session 1*. Introduction.** The benefits to child development of book-sharing are explained, and the importance of establishing a book-sharing routine. Basic principles of dialogic reading are outlined, including following the child's lead and being positive, as well as techniques such as pointing and naming, asking 'who/what/where' as well as open ended questions to engage the child, linking the book content to the child's own experience, and generally encouraging a dialogue.
Book of the week 'Handa's Surprise' by Eileen Browne.
***Session 2*. Elaborating and Linking**. Picking up on the child's focus of interest and elaborating on it. Making links between the book content and the child's own experience.
Book of the week 'Little Helpers' by Lynne Murray and Peter Cooper
***Session 3*. Numbers and Comparisons** (including Executive function skills). Practising activities that promote enumerating and making comparisons, working memory, inhibition and shifting.
Book of the week 'Handa's Hen' by Eileen Browne
***Session 4*. Emotions.** Talking about the feelings of the book characters. Naming feelings and contextualizing them. Linking the book characters' feelings to the child's own emotional experience.
Book of the week 'Hug' by Jez Alborourgh
***Session 5*. Intentions**. Discussing book characters' desires, intentions and beliefs and why they might behave as they do.
Book of the week 'Harry the Dirty Dog' by Gene Zimmerman
***Session 6.* Perspectives.** Highlighting the fact that different characters in the book can have differing desires, intentions, beliefs and perspectives, and how these influence the characters' behaviour.
Book of the week 'Harry by the Sea' by Gene Zimmerman
***Session 7*. Relationships, and Summary and taking forward**. Discussing everyday family relationships, including conflicts and their resolution. Summary of the key learning points from sessions 1-6; and discussion of continuing with regular book-sharing (including accessing books).
Book of the week 'The wrong side of the bed' by Edward Ardizzone

Figure 1 Two sample Powerpoint slides from Session 4 on Feelings



Sample size

The required study sample size was calculated on the basis of an earlier meta-analysis of book-sharing interventions (Dowdall, 2015), where three trials (two US, one South Africa) using a similar format to that used in the current study showed the average effect size on language development (one of our main primary outcomes) 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 and Control 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.

Recruitment

Parents (or other carers such as grandparents) attending the 12 study Children's Centres with a child aged between 28 and 45 months old, and who regularly spoke English at home, were approached directly by the Trial Manager and invited to join the study. If parents were not present on recruitment days, they were invited on another occasion by Children's Centre staff to provide contact information for the study team. Children with a significant developmental disorder (e.g., autism, Down's syndrome) were excluded. Those in the Intervention group centres were told that we were conducting a study about possible effects on child development of additional input provided within Children's Centre, and that participation would involve three assessments of their child's development, spaced over a 10-month period, and attendance at seven, weekly book-sharing sessions in the Children's Centre. Those in the Control group Centres were told that we were interested in child development and the input provided within Children's Centres, and that their participation would involve three assessments of their child's development over the same, 10-month period. All participants were told that they would receive gratuities for assessment sessions. 108 participants were recruited into the Control group, and 110 into the Intervention group. One additional participant was originally assigned to the Intervention group but they were subsequently removed from all analyses due to a diagnosis that made them ineligible for the study.

Of 343 parents who were contacted about the study, 96 (28%) either declined to participate or did not respond. Numbers of decliners/non-responders were 61 out of 190 (32%) of the target Intervention group, and 35 out of 153 (23%) of the target Control group. (An additional 28 (8%) of those initially approached were found to be ineligible.)

Assessments

Child assessments comprised a range of directly administered developmental tasks, and parent reports of child development. Parenting assessments comprised direct observations during parent-child interactions and self-report questionnaires. A standard proforma was used for parents to provide demographic information.

Our child assessments were grouped into a number of over-arching, composite, developmental areas, namely, Cognitive Development (comprising Language, Attention and Executive Function), Social Development, Behaviour Problems, and Emotion Regulation. The three dimensions of Cognitive Development were our primary outcomes. In order to achieve a robust measure for each area of development, we used a number of component assessments, balancing their number with both the importance of the outcome in question to the study, and the ability of small children to undertake multiple assessments. The assessments of parenting were, given the nature of the intervention, focused on book-sharing; for these interactions we scored dimensions known to be of particular benefit to child cognitive and social development. These included measures of how sensitive parents were to the child's cues and interests, how much they provided support and stimulation for the child's thinking ('cognitive scaffolding'), how much they talked about the book characters' mental states, and the extent to which they had a reciprocal, two-way dialogue with the child. As we were interested in whether an intervention concerning book-sharing might also be of benefit to parenting in other contexts, and particularly those relevant to children's behaviour problems and emotion regulation, we also assessed parenting during a task presenting a challenge to the child – not to touch attractive toys for a few minutes. The study measures are set out in Table 2, and the nature of the variables arising from these assessments that were analysed are shown in Appendix 1.

Assessments were conducted at Baseline (before the Intervention), Post intervention, and at Follow-up approximately 4- 6 months after the end of the intervention. We considered the Post-intervention assessment as the principal measure of parenting outcome, but the Follow-up assessment our principal measure of child development outcome (because the time for children to have benefitted from the intervention at the point of the immediate Post treatment assessment was deemed too short). Assessments were carried out by trained researchers who were blind to study group; parents were asked not to reveal their group allocation to researchers. Child development and parenting assessments were videotaped and scored by other trained researchers who were blind to study group.

Table 2 Measures and timing of assessments

			Baseline	Post-	Follow-up
CHILD					
Cognitive	Language	EYT Expressive	x	x	X
		CELF Receptive			X
	Attention composite	ECVT	x	x	X
		EYT Go-No-go *(consistency)*	x	x	X
		Three Toy Play Task *(persistence and quality)*	x	x	X
		Parent Report *(SDQ/CSBQ)*	x	x	X
	Executive Function composite	WPPSI Block Design		x	X
		EYT Go-No-go *(inhibition)*	x	x	X
		EYT Card Sort *(shifting)*			X
		Digit Span *(working memory)*		x	X
		Following Instructions *(working memory)*		x	X
		LAB-TAB *(persistence)*	x		X

		Parent Report (CSBQ) *(Self-regulation)*	x	x	X
Social Development	Social Development composite	Theory of Mind tasks	x	x	X
		Expressive Emotion understanding	x	x	X
		Help task (*empathy*)	x	x	X
		Altruism task			X
		Parent report (SDQ)	x	x	X
Behaviour Problems		Don't touch task *(defiance)*	x		X
		Parent report (SDQ)	x	x	X
Emotion regulation		Parent report (CSBQ)	x	x	X
PARENTING					
Book-sharing		Sensitivity	x	x	
		Reciprocity (Scale)	x	x	
		Reciprocity (Count)	x	x	
		Cognitive Scaffolding	x	x	
		Mental State Talk	x	x	
Behaviour Management		Don't touch *(physical and verbal guidance)*	x	x	
		Parent report (DVQ) *(harsh psychological, non-harsh negative, positive)*	x	x	X

A few measures that were administered are not shown, as distributions of scores were problematic (e.g., because a great number of participants had optimum scores at baseline, and so had no room for improvement), and were not included in the final analyses.

Hypotheses

We hypothesised that, at Post and Follow-up assessments, both child development and parenting would be better in the Intervention vs. the Control group. As well as assessing the direct effects of the intervention, we investigated whether the effects varied depending on factors known to be associated with child development (i.e., were moderated by them), although no specific hypotheses concerning such effects were made. These factors comprised: parental income and education, multilingualism status, child ethnic group, parental mental health [HADS] and stress [PSI] (both assessed by questionnaire), child

gender, age, birth order, presence of siblings and presence of child-age siblings, and baseline parenting (i.e., the measure of Sensitivity assessed during book-sharing). Finally, we predicted that any benefits of the intervention to child development would be brought about, or mediated, by improvements in parenting.

Statistical analyses

Our primary analyses were conducted according to 'Intention To Treat' (ITT), that is, including *all* those recruited into the Intervention group, regardless of whether or not they attended, or completed the minimum target number of sessions. However, because this was an efficacy trial (that is, it was not a trial conducted under normal 'real-world' conditions - for example, the intervention was not delivered by Children's Centre staff known to the parents, but by Research Facilitators), we also conducted analyses concerning the Control group and the Intervention group participants who actually undertook the intervention as intended (attending at least 5 of the 7 sessions) - that is, a 'Per-protocol' analysis.

A further, exploratory, post-hoc, analysis was made of the primary study outcomes, with a view to future possible wider implementation. Thus, we were concerned to establish whether parental *engagement* in the intervention was important for child outcome, over and above simple attendance. Accordingly, after each training session, facilitators made ratings concerning the extent of participants' active engagement in the session and the value participants attached to the book-sharing; these were used to define an engaged subgroup of those who received the intervention, and these participants were compared to Control group participants.

Statistical methods

Before the statistical analyses of the effects of the intervention were undertaken, scores on the different measures were examined. A few measures were found not to be suitable for analysis (e.g., because a great number of participants had optimum scores at baseline, and so had no

room for improvement[3]). In addition, scores on the individual components of the main child development outcomes were checked (i.e., their distributions and their correlations with the other components). In each case (Attention, Executive Function, Social Development), a satisfactory composite measure was derived by combining the individual components by using standardised scores or factor analysis, as appropriate. Analyses of individual components were considered secondary analyses[4].

The comparison of the Intervention and Control groups on the outcomes at each assessment were assessed using linear mixed multilevel models, where clustering within Children's Centre and, where applicable, repeated measures (that is, outcomes measured at both Post and Follow-up assessments), were accounted for. Since child gender and age, parent education, and family multilingualism, and are all known to have important effects on child development, these factors were taken into account (co-varied) when conducting the analyses of the effects of intervention, as was the amount of time parents read to the child at Baseline where this measure made an independent contribution to child performance on any primary outcome. In addition, the child's performance at the Baseline assessment (where available) was also taken into account as a covariate. For some measures, no baseline assessment was available because the age range of the children made administration inappropriate (Receptive Language, and components of Executive Function (e.g., 'Following Instructions')). For these measures, the child's Expressive Language score at Baseline was taken into account as a covariate instead. Multiple imputation was used for any missing data on only the primary Cognitive Development outcomes and the key secondary outcome of child Social Development using the Intention to Treat population. For all other analyses, only available

[3] These were the Receptive measure from the Emotional Understanding Task, Lab-tab anger and distress, the harsh physical measure from the Don't touch task, and the Violent Discipline subscale from the parent self-report Violence and Discipline questionnaire. The language development parent report (CDI) was not a suitable measure for analyses due to ceiling effects

[4] See Appendix 1 for details of variables analysed

data were included. Statistical tests were two-sided, and performed using a 5% significance level. No adjustments were made for multiple comparisons. This is because the primary endpoints are associated with each other and an adjustment would have over-corrected (Schulz 2005). Since the number of primary, secondary and exploratory endpoints and related analyses was large, care must be taken when considering the findings as significant results may be due to chance. Transformations, and other models (e.g. negative binomial) were used when distributions required them. Adjusted means and 95% confidence intervals (CI) are derived from the models using the observed margins for categorical covariates, and the average age at each visit for each endpoint. Standardised effect sizes that accounted for clusters and intra-cluster correlations (ICC) were produced for the great majority of outcomes. Moderation analyses were performed using linear mixed models as described for the primary analyses, but also included the relevant moderator by treatment interaction and moderator by treatment by visit interaction where appropriate. Mediation analyses were conducted using structural equation modelling.

Results

Sample

The large majority of those who were recruited into the Intervention group - 94 out of 110 (85%) - completed the intervention as intended (attending at least five of the seven 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). Of those completing the programme, 69% were rated as having engaged well. The same Control group was used in all analyses.

The key demographic characteristics at baseline of the different groups are shown in Table 3. In the Intervention group, 101 (92%) and 97 (88%) families were seen at Post and Follow-up assessments, respectively, and in the Control group the numbers 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%).

Primary Analyses

Control vs. Intention to treat (ITT) Intervention group analyses

No formal comparisons were made of the demographic characteristics of the Control vs. Intervention group participants, given the random allocation of groups. Nevertheless, it can be seen in Table 3 that multilingualism and non-White ethnic status were somewhat more common in the (ITT) Intervention group than in the Controls, and Intervention group parents read less to their child at baseline.

Results of Control vs. ITT Intervention group analyses at Post and Follow–up assessments are shown in Table 4a and Appendix 2. Effects of the intervention on child development were not statistically significant, although, 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)) $p = 0.13$). 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 0.38 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.

Parenting during book-sharing at the Post intervention assessment showed statistically significant benefits of the intervention on all measures, with the largest effects observed on

the Sensitivity and Cognitive Scaffolding measures. By contrast, parenting during the Don't Touch discipline task showed no differences between groups.

There was no evidence that the improvements in parenting during book-sharing indirectly benefitted children's development.

Control vs. Per protocol intervention group analysis

As noted when discussing the sample, only 16 of those recruited into the Intervention group failed to attend the requisite number of sessions and, accordingly, the demographic characteristics of the Per protocol population were broadly similar to those for the full ITT Intervention group, having a somewhat greater percentage of non-White, multilingual families than in the Controls. Despite the small change in numbers, in general, benefits of the intervention to child development were more evident when considering the Per protocol group (see Table 4b and Appendix 3). Thus, 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)), $p = 0.27$), and attention ES = 0.34 (95% CI (-0.02, 0.70)), $p = 0.06$).

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 0.48 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 0.33 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 during book-sharing showed similar statistically significant positive effects as for the ITT analysis on all measures, with the largest effects observed on the

Sensitivity and Cognitive Scaffolding measures; and those for parenting discipline again showed no significant effects of the intervention.

Secondary and exploratory analyses

Engagement in the intervention

As noted previously, of those who *did* complete the programme, 69% were rated as having engaged well. While the demographic characteristics of those who engaged with the intervention were broadly similar to those for Control group participants, it was notable that those who did not engage well, and particularly those who did not attend the requisite number of sessions (the non Per-protocol population) appeared to be at somewhat higher risk than those who did engage in terms of low income and education (see Table 3).

Given that analyses concerning the Engaged Intervention group vs. Controls were conducted as post hoc exploratory analyses, we considered only the four main primary child outcomes (i.e. Expressive Language, Comprehension, the Attention composite and the Executive Function composite). These are shown in Table 5, where it can be seen that effect sizes comparing the Engaged Intervention group to Control were larger than those seen for the ITT and Per Protocol populations.Thus, 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).

Table 3 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	17.6	16.4	10.6	50.0	7.7	28.9
</= A Level	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	n = 98	n = 97	n = 87	n = 10	n = 61	n = 36
(mins/week) M (SD)	114.6(70.71)	71.2(63.89)	72.0(64.34)	64.2(62.61)	76.6(69.30)	61.9(53.12)

Table 4.

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

		a) Intervention vs Control (ITT)		b) Intervention vs Control (Per Protocol)	
Variables	Time point	Standardized effect size (I-C) (95% CI) [a]	*P*	Standardized effect size (I-C) (95% CI) [a]	*P*
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 individual components of compositive variables and parental discipline in Appendix 2 for ITT population and Appendix 3 for per protocol population. Multiple imputation was used for missing data for the primary outcomes and social development composite for the ITT analyses only. [a]Adjusting 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. *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 5

Adjusted mean scores and intervention effects for the different study populations (ITT, Per-protocol, Engaged) for Primary child outcomes

Variable	Time point	Intervention n	Control N	Intervention adjusted mean (95% CI)	Control adjusted mean (95% CI)	Adjusted difference (95% CI)	Standardized effect size (95% CI)	*P* value
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)	-0.04 (-0.40, 0.32)	0.81
	Follow - up	110	108	28.13 (26.79, 29.47)	26.72 (25.42, 28.01)	1.41 (-0.49, 3.31)	0.34 (-0.12, 0.80)	0.13
Comprehension CELF*	Follow- up	110	108	56.94 (54.21, 59.66)	56.15 (53.48, 58.81)	0.79 (-3.21, 4.79)	0.07 (-0.28, 0.43)	0.69
Attention composite	Post	110	108	0.20 (0.07, 0.34)	0.21 (0.10, 0.33)	-0.01 (-0.19, 0.16)	-0.02 (-0.39, 0.34)	0.89
	Follow- up	110	108	0.41 (0.27, 0.54)	0.35 (0.22, 0.47)	0.06 (-0.12, 0.24)	0.11 (-0.24, 0.47)	0.53
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)	-0.19 (-0.69, 0.31)	0.41
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)	-0.02 (-0.37, 0.33)	0.91
	Follow - up	85	98	29.25 (28.02, 30.48)	27.42 (26.21, 28.63)	1.83 (0.06, 3.59)	0.46 (0.02, 0.91)	0.04
Comprehension CELF*	Follow- up	83	100	59.99 (56.58, 63.40)	57.56 (54.10, 61.02)	2.43 (-2.45, 7.31)	0.22 (-0.22, 0.66)	0.29
Attention composite	Post	58	79	0.22 (0.11, 0.33)	0.22 (0.12, 0.31)	0.00 (-0.14, 0.15)	0.01 (-0.38, 0.40)	0.96
	Follow- up	56	78	0.45 (0.35, 0.56)	0.32 (0.23, 0.41)	0.13 (-0.01, 0.28)	0.34 (-0.02, 0.70)	0.06

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)	-0.01 (-0.53, 0.52)	0.98
Engaged group								
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, and for EYT. [#]adjusting for base-line book-sharing. *adjusting for baseline EYT

Discussion

The intervention aimed to benefit child development by improving parents' book-sharing. Such improvement was achieved across all the dimensions of book-sharing we measured, and was evident in both Per-Protocol and Intention to Treat analyses. Indeed, the intervention effects were large for both parental Sensitivity and Cognitive scaffolding. While, in the main, the primary outcomes concerning child cognitive development did not show statistically significant effects of the intervention, the magnitude of benefit to language was generally consistent with effects found in previous trials. 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-behaviour difficulties.

Two aspects of our findings in particular require further consideration. First, given the substantial impact of the intervention on parental behaviour during book-sharing, the question arises as to why there were not greater benefits to child development. There are a number of possible explanations. First, it is possible that the 4-6 month period of follow-up was too short for changes in parenting to have had a positive impact on the children, and that more positive effects will emerge in due course. It is of relevance to note that examination of the results at Post and Follow-up showed there was a trend for larger effects sizes to be seen at Follow up for dimensions of child development measured at both points, consistent with the possibility that the performance of children in the intervention group may have been on a positive trajectory. This possibility is currently being examined in a further follow-up of the study population.

A second possible explanation for any lack of benefit of the intervention to child outcome may be that the children in our sample were already receiving input through Children's Centres' routine programmes (or other pre-school experiences) that was of sufficient quantity and quality to support their development. Similarly, national programmes to inform parents about the benefits of reading to children, and the regular provision to families of sets of children's books from early infancy, may have meant that, although significant gains in the quality of parental book-sharing could be achieved with the intervention, the scope for effecting additional change in child development may have been limited.

With regard to child social and emotional-behavioural outcomes, although the content of our intervention and choice of books was designed to elicit particular kinds of positive parent-child interaction that were hypothesised to enhance these areas of child development, it did not contain the explicit elements of practice in specific interactive strategies relevant to these outcomes that have been used in some other, more successful,

interventions for these child outcomes using a general book-sharing format (e.g. Howard et al., 2016; Chacko et al., 2018; Lecce et al., 2014). By the same token, the fact that parenting during book-sharing showed substantial benefits of the intervention, while parental discipline in the context of a challenge task showed none, similarly suggests that for the latter parenting outcomes to benefit, more specific focused strategies need to be incorporated into DBS, including those concerning negative parenting practices, if change in these behaviours is to be effected.

A further consideration relevant to interpreting our results for child development concerns the characteristics of the intervention and control group samples. As noted earlier, demographic characteristics differed somewhat between groups. We also found on inspection of baseline data that the control group generally performed better than the intervention sample in terms of both the quantity and quality of parents' book-sharing and child performance (see Appendix 4 for these raw Baseline scores). To mitigate the effects of these differences, our analyses controlled for key demographic characteristics and baseline child functioning. Nevertheless, it remains possible that this did not entirely compensate for any initial disadvantage in the intervention group. Given evidence that achievement gaps between well and more poorly performing children generally increase over time (e.g., Fernald, Marchman, & Weisleder, 2013; ASPE, 2014), in the absence of an intervention effect one might have expected the performance of intervention group children to show a relative *decline* over time. In fact, the opposite was the case. One possible conclusion, therefore, is that the intervention played a buffering role, by countering the expected widening achievement gap over time.

The fact that we principally showed benefits to child development where parents engaged effectively with the intervention highlights a second key issue to arise from our study. A lower level of educational achievement characterised those who, even though

participating in the intervention, were judged not to have engaged (29% had received only GCSE, or no educational qualification, in the disengaged group, vs. only 7% of those who did engage). It was also the case that being on a very low income (under £16,000) was twice as common in those who did not engage (33%) vs. those who did (15%). Finally, there was some suggestion that Asian parents were less likely to engage with the intervention (27% non-engaged vs. 15 % engaged). These possible differences need to be confirmed in further research; however, if confirmed, the issue of how better to frame book-sharing interventions so as to be able to engage particular groups is one that is important to address. Our previous work in training parents in book-sharing has been conducted in peri-urban settlement communities in South Africa, where extreme poverty, and limited education are common (Cooper et al., 2014; Vally et al. 2016), and it was striking, in those populations, to find the level of parental motivation to engage with the intervention to be very high. Given that this intervention in the UK was closely modelled on that successfully delivered in South Africa, low education and poverty do not, therefore, per se, appear to explain the greater failure to engage on the part of some UK parents. Rather, it would seem important to understand more about the significance of the intervention in different groups. In our South Africa work, the idea that a child could be helped to become better equipped to learn and to be well-prepared for school through parental book-sharing was highly motivating for parents who considered school and educational success to be a route out of inter-generational poverty. Such motivations may be far 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, or that what they do as parents will not stand to make a difference. A greater understanding of parental perceptions may therefore be required in order to deliver interventions effectively.

Finally, a further factor that might have limited engagement in the current efficacy study is that the intervention was delivered by research facilitators who were unfamiliar to the parents, and who delivered the program in isolation from other centre support. There is good evidence that the quality of relationship with intervenors is important in influencing parents' attendance and engagement (e.g., Robbins, Turner, Alexander, & Perez, 2003). Thus, future work may achieve greater engagement by integrating interventions into centres' wider work with families so that they are delivered by staff who have already built relationships with parents.

References

Allen G. (2011) *Early Intervention: the Next Steps*, London: *Cabinet Office*, p23).

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

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

Burgoyne K, Gardner R, Whiteley H, Snowling MJ, 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.

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 : the official journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *47*(1), 79–93. https://doi.org/10.1080/15374416.2016.1266648J

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.

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, p3.

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 GJ, Murray L, Gardner F, Hartford L, Cooper PJ (2019) Shared Picture Book Reading Interventions for Child Language Development: A Systematic Review and Meta-Analysis *Child Development* doi: 10.1111/cdev.13225.

Fernald, A., Marchman, V.A., Weisleder, A. (2013) SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science, 16* (2), 234-248.

Fletcher, K.L., Reese, E. (2005) Picture book reading with young children: A conceptual framework *Developmental Review, 25* (1) 64-103

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, 15:1, 24-38, DOI: 10.1080/15295192.2015.992736

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society*, *11*, 49–76.

Howard, S., Powell, T., Vasseleu, E., Johnstone, S., Melhuish, E. (2016) Enhancing preschoolers' executive functions through embedding cognitive activities in shared book reading *Educ Psychol Rev 28*(1) DOI 10.1007/s10648-016-9364-4

Law, J., Charlton, J., McKean, C., Beyer, F., Fernandez-Garcia, C., Mashayekhi, A., Rush, R. (2018) Nuffield Foundation FINAL REPORT Parent-child reading to improve language development and school readiness: A systematic review and meta-analysis https://www.nuffieldfoundation.org/news/reading-pre-school-children-boosts-language-skills-eight-months

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

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**.**

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.

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(12):1370-1379 DOI: 10.1111/jcpp.12605

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

Peskin, J, Astington, J. (2004) The effects of adding metacognitive language to story texts *Cognitive Development 19(2)*. 253-273

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.

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.

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(3)*. 189-201.

Schulz KF, Grimes DA. Multiplicity in randomised trials I: endpoints and treatments. *Lancet* 2005; 365: 1591–95.

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language, 24,* 123–138.

Sylva, K. (2014). The role of families and pre-school in educational disadvantage. *Oxford Review of Education, 40,* 680-695

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.

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.

Appendix 1 Variables derived from assessments for statistical analysis

CHILD			
Cognitive	Language	EYT Expressive	Total score
		CELF Receptive	Mean % score on 3 subscales- Sentence structure, Concepts and Following Directions, Basic Concepts
	Attention composite		Mean of z scores for individual measures
	Attention individual measures	ECVT	% time attending to target
		EYT Go-No-go *Consistency*	Standard Deviation of time to respond in correct Go trials
		Three Toy Play Task *persistence* *quality*	Longest play bout on a single toy Mean quality on 4–point scale, scored every 30 seconds
		Parent Report *(SDQ/CSBQ)*	Total of SDQ Q15 (reversed) and Q25, and CBSQ Q4 (reversed)
	Executive Function composite		Factor analysis score derived from individual measures
	Executive Function individual measures	WPPSI Block Design	Total score
		EYT Go-No-go *(inhibition)*	Proportion correct No-Go trials minus Proportion of incorrect Go trials
		EYT Card Sort *(shifting)*	Total score
		Digit Span *(working memory)*	Total trials score
		Following Instructions *(working memory)*	Total accuracy score

		LAB-TAB *(persistence)*	Mean of 5–point persistence quality scale scores, scored every 20 seconds
		Parent Report (CSBQ) *(Self-regulation)*	Total of Behavioural and Cognitive Regulation subscale scores
Social Development	Social Development composite		Factor analysis score derived from individual measures
	Social Development individual measures	Theory of Mind tasks	Pass score for 9 tasks
		Expressive Emotion understanding	Total emotion faces named correctly
		Help task (*empathy*)	Degree x swiftness of help given
		Altruism task	Total number of sweets given
		Parent report (SDQ)	Pro-social scale score
Behaviour Problems		Don't touch task *(defiance)*	Mean score on 6-point defiance scale, scored every 20 seconds
		Parent report (SDQ)	Sum of Emotional, Hyperactivity, Conduct and Peer problems subscales
Emotion regulation		Parent report (CSBQ)	Total of Qs 1 (reversed), 6 (reversed), 7, 10, 13, 14
PARENTING			
Book-sharing		Sensitivity	Score on 5-point scale
		Reciprocity (Scale)	Score on 5-point scale
		Reciprocity (Count)	Average of % times parents responds to child and % times child responds to parent
		Cognitive Scaffolding	Mean of z scores for: % parent responses that are cognitively enriching, number of elicitations, number of book elements identified, number of grammatically complex elements

		Mental State Talk	Number of Mental State terms used in non-simple constructions
Behaviour Management		Don't touch *physical guidance* *verbal guidance*	Mean of scores on 4-point scale scored every 20 seconds for both measures
		Parent report *(harsh psychological, non-harsh negative, positive)* (DVQ)	Total of each subscale score

Appendix 2. Summary of the results of the statistical analyses of the Individual Components of composite child variables and Parent discipline for Intention to Treat population.

	Visit		Intervention n	Control n	Intervention adjusted mean (95% CI)	Control adjusted mean (95% CI)	Adjusted difference (95% CI)	Standardised effect Size (Intervention-Control) (95% CI)	P-value
INDIVIDUAL COMPONENTS OF COMPOSITE CHILD VARIABLES									
ECVT (%)	Post		94	98	58.996 (55.133, 62.858)	67.430 (63.644, 71.215)	-8.434 (-13.948, -2.920)	-0.46 (-0.75, -0.16)	0.0029
	Follow-up		90	94	66.964 (62.932, 70.996)	63.096 (59.147, 67.045)	3.868 (-1.849, 9.584)	0.22 (-0.10, 0.53)	0.1836
Play persistence	Post		79	91	113.376 (105.517, 121.235)	105.977 (98.675, 113.279)	7.398 (-3.493, 18.290)	0.25 (-0.12, 0.61)	0.1818
	Follow-up		75	90	113.005 (105.800, 120.209)	115.628 (109.052, 122.203)	-2.623 (-12.488, 7.242)	-0.08 (-0.37, 0.22)	0.6003
Play Quality	Post		79	91	2.815 (2.709, 2.921)	2.807 (2.709, 2.905)	0.008 (-0.139, 0.156)	0.02 (-0.29, 0.33)	0.9104
	Follow-up		75	90	3.025 (2.910, 3.140)	3.011 (2.905, 3.117)	0.013 (-0.146, 0.173)	0.03 (-0.33, 0.39)	0.8689
Task consistency (low good)	Post		88	91	0.230 (0.211, 0.249)	0.230 (0.211, 0.248)	-0.000 (-0.027, 0.027)	-0.00 (-0.39, 0.39)	0.9945
	Follow-up		84	90	0.196 (0.181, 0.212)	0.198 (0.183, 0.213)	-0.002 (-0.024,0.020)	-0.02 (-0.28,0.23)	0.8581
Parent report attention	Post		101	103	3.675 (3.465, 3.886)	3.502 (3.294, 3.710)	0.173 (-0.130, 0.476)	0.15 (-0.11, 0.41)	0.2615
	Follow-up		96	100	3.766 (3.521, 4.010)	3.574 (3.334, 3.813)	0.192 (-0.154, 0.538)	0.19 (-0.15, 0.53)	0.2753

Block Design*	Post		100	103	13.03 (12.22, 13.83)	13.20 (12.39, 14.01)	-0.18 (-1.34, 0.99)	-0.06 (-0.45, 0.33)	0.7573
	Follow-up		95	100	15.38 (14.65, 16.11)	16.42 (15.68, 17.16)	-1.04 (-2.10, 0.02)	-0.30 (-0.61, 0.01)	0.0550
Card Sort*	Follow-up		90	100	11.32 (10.18, 12.46)	10.25 (9.11, 11.40)	1.06 (-0.57, 2.70)	0.31 (-0.17, 0.79)	0.1810
Following Instructions*	Post		98	103	49.09 (44.01, 54.16)	47.72 (42.79, 52.66)	1.36 (-5.93, 8.66)	0.06 (-0.24, 0.35)	0.7126
	Follow-up		93	100	59.66 (54.32, 65.01)	55.89 (50.69, 61.08)	3.77 (-3.92, 11.46)	0.16 (-0.16, 0.48)	0.3345
Digit span*	Post		100	102	12.59 (11.51, 13.67)	13.92 (12.85, 14.98)	-1.33 (-2.89, 0.23)	-0.31 (-0.67, 0.05)	0.0949
	Follow-up		96	100	17.40 (16.47, 18.33)	17.13 (16.22, 18.04)	0.27 (-1.08, 1.62)	0.05 (-0.21, 0.32)	0.6946
Impulse control	Post		95	95	0.23 (0.16, 0.30)	0.26 (0.19, 0.33)	-0.03 (-0.13, 0.07)	-0.11 (-0.45, 0.24)	0.5221
	Follow-up		89	91	0.39 (0.31, 0.47)	0.43 (0.35, 0.51)	-0.04 (-0.15, 0.07)	-0.17 (-0.64, 0.30)	0.4666
Persistence Quality	Follow-up		87	99	3.44 (3.31, 3.57)	3.47 (3.36, 3.59)	-0.04 (-0.21, 0.14)	-0.06 (-0.38, 0.25)	0.6885
Parent report (CSBQ)	Post		102	103	14.85 (14.24, 15.46)	14.77 (14.14, 15.40)	0.08 (-0.81, 0.97)	0.03 (-0.27,0.33)	0.8425
	Follow-up		96	100	15.41 (14.74, 16.09)	15.10 (14.41, 15.78)	0.31 (-0.65, 1.28)	0.13 (-0.27,0.52)	0.5064
Expressive (Emotion Understanding)	Post		100	100	3.55 (3.13, 3.97)	3.70 (3.27, 4.13)	-0.15 (-0.76, 0.46)	-0.11 (-0.56, 0.34)	0.6209
	Follow-up		97	99	4.71 (4.35, 5.08)	5.05 (4.68, 5.42)	-0.33 (-0.86, 0.19)	-0.19 (-0.48, 0.11)	0.1962
Helpfulness (Empathy)	Post		92	98	6.72 (5.74, 7.70)	7.87 (6.88, 8.85)	-1.15 (-2.56, 0.26)	-0.31 (-0.68, 0.07)	0.1052
	Follow-up		87	97	8.29 (7.34, 9.24)	8.19 (7.25, 9.13)	0.10 (-1.24, 1.45)	0.03 (-0.31, 0.36)	0.8729

Altruism	Follow-up		93	97	3.48 (3.03, 3.93)	3.67 (3.23, 4.11)	-0.19 (-0.83, 0.46)	-0.09 (-0.40, 0.22)	0.5704
Theory of Mind	Post		99	100	0.40 (0.36, 0.43)	0.38 (0.35, 0.42)	0.02 (-0.03, 0.06)	0.09 (-0.19, 0.38)	0.5271
	Follow-up		94	99	0.49 (0.45, 0.52)	0.47 (0.43, 0.50)	0.02 (-0.04, 0.07)	0.11 (-0.23, 0.45)	0.5134
Prosocial (SDQ)	Post		101	103	7.55 (7.23, 7.87)	7.41 (7.08, 7.74)	0.15 (-0.32, 0.61)	0.10 (-0.21, 0.41)	0.5068
	Follow-up		96	100	7.86 (7.52, 8.21)	7.39 (7.05, 7.74)	0.47 (-0.02, 0.96)	0.35 (-0.02, 0.71)	0.0601
PARENT DISCIPLINE							**Adjusted difference (95% CI)**	**Standardised effect size (95% CI)**	
Harsh Psychological (parent report)(low good)	Post		101	103	1.70 (1.52, 1.89)	1.48 (1.29, 1.66)	0.23 (-0.04, 0.50)	0.25 (-0.05, 0.55)	0.1000
	Follow-up		97	100	1.52 (1.33, 1.71)	1.57 (1.38, 1.75)	-0.05 (-0.32, 0.22)	-0.05 (-0.35, 0.25)	0.7320
Non-Harsh – Negative (parent report)	Post		101	103	5.47 (5.11, 5.83)	5.22 (4.85, 5.59)	0.25 (-0.27, 0.77)	0.15 (-0.17, 0.47)	0.3196
	Follow-up		97	100	5.35 (4.96, 5.73)	5.10 (4.70, 5.49)	0.25 (-0.31, 0.80)	0.17 (-0.21, 0.55)	0.3588
Non-Harsh – Positive (parent report)	Post		101	103	4.77 (4.59, 4.95)	4.58 (4.40, 4.76)	0.19 (-0.07, 0.45)	0.26 (-0.10, 0.61)	0.1545
	Follow-up		97	100	4.81 (4.65, 4.96)	4.80 (4.65, 4.95)	0.01 (-0.21, 0.22)	0.01 (-0.24, 0.26)	0.9620
Physical Guidance ^	Post		100	102	0.50 (0.36, 0.64)	0.45 (0.31, 0.59)	0.05 (-0.15,0.25)	0.13 (-0.39, 0.64)	0.5897

Verbal guidance	Post		100	102	1.58 (1.36, 1.81)	1.60 (1.37, 1.84)	-0.02 (-0.35, 0.31)	-0.03 (-0.46, 0.40)	0.882 5

Adjusting for available baselines (and baseline by visit interaction where appropriate), child's age at each visit, child's gender, multilingualism and parental education. *adjusting for EYT baseline ^square root transformation used. [+] Follow-up visit only used in subjects where there was no video of book-sharing at post (3 participants). Intra-cluster correlation (ICC) was between 0 and 0.07 for all measures. For measures where a low value is a good, a negative standardised effect size is considered a positive effect of the intervention. Multiple imputation was used for missing data in the analyses of the primary outcomes and the child social development composite.

Appendix 3. Summary of the results of the statistical analyses of the Individual components of child composite variables and Parent discipline for the Per Protocol population.

	Visit	Intervention n	Control n	Intervention adjusted mean (95% CI)	Control adjusted mean (95% CI)	Adjusted difference (95% CI)	Standardised effect size (Intervention-control) (95% CI)	P-Value
INDIVIDUAL COMPONENTS OF COMPOSITE CHILD VARIABLES								
ECVT (%)	Post	86	98	58.602 (54.650, 62.553)	67.820 (64.115, 71.526)	-9.216 (-14.747, -3.691)	-0.49 (-0.79, -0.20)	0.0012
	Follow-up	81	94	67.512 (63.204, 71.821)	63.437 (59.446, 67.428)	4.076 (-1.870, 10.021)	0.23 (-0.11, 0.57)	0.1779
Play persistence	Post	70	91	111.692 (103.364, 120.019)	106.149 (98.868, 113.430)	5.543 (-5.688, 16.774)	0.19 (-0.19, 0.56)	0.3313
	Follow-up	65	90	114.281 (106.543, 122.018)	115.766 (109.209, 122.322)	-1.485 (-11.743, 8.773)	-0.04 (-0.35, 0.26)	0.7752
Play Quality	Post	70	91	2.830 (2.717, 2.944)	2.805 (2.706, 2.903)	0.026 (-0.128, 0.180)	0.05 (-0.27, 0.37)	0.7399
	Follow-up	65	90	3.056 (2.930, 3.182)	3.011 (2.903, 3.118)	0.045 (-0.123, 0.214)	0.10 (-0.28, 0.48)	0.5957
Task consistency (low good)	Post	80	91	0.226 (0.207, 0.245)	0.228 (0.210, 0.247)	-0.002 (-0.030, 0.025)	-0.04 (-0.43, 0.36)	0.8607
	Follow-up	76	90	0.193 (0.177, 0.210)	0.197 (0.182, 0.212)	-0.004 (-0.027, 0.018)	-0.05 (-0.32, 0.22)	0.7099

Parent report attention	Post	91	103	3.717 (3.495, 3.939)	3.514 (3.306, 3.723)	0.202 (-0.109, 0.514)	0.18 (-0.10, 0.45)	0.2020
	Follow-up	86	100	3.863 (3.608, 4.117)	3.585 (3.349, 3.821)	0.278 (-0.073, 0.628)	0.27 (-0.07, 0.62)	0.1201
Block Design*	Post	90	103	13.08 (12.25, 13.91)	13.27 (12.48, 14.07)	-0.19 (-1.37, 0.98)	-0.07 (-0.49, 0.35)	0.7374
	Follow-up	85	100	15.77 (15.05, 16.49)	16.47 (15.77, 17.17)	-0.70 (-1.72, 0.33)	-0.20 (-0.49, 0.09)	0.1704
Card Sort*	Follow-up	81	100	11.32 (10.14, 12.49)	10.30 (9.15, 11.45)	1.02 (-0.64, 2.68)	0.30 (-0.19, 0.78)	0.2050
Following Instructions*	Post	88	103	49.58 (44.21, 54.95)	48.37 (43.43, 53.32)	1.20 (-6.30, 8.71)	0.05 (-0.26, 0.35)	0.7519
	Follow-up	83	100	60.43 (54.81, 66.04)	56.39 (51.25, 61.53)	4.04 (-3.79, 11.87)	0.17 (-0.16, 0.50)	0.3103
Digit span*	Post	90	102	12.92 (11.78, 14.06)	14.05 (12.99, 15.11)	-1.13 (-2.73, 0.47)	-0.27 (-0.65, 0.11)	0.1639
	Follow-up	86	100	17.92 (16.96, 18.88)	17.25 (16.36, 18.14)	0.68 (-0.68, 2.03)	0.13 (-0.13, 0.39)	0.3248
Impulse control	Post	87	95	0.23 (0.16, 0.30)	0.26 (0.19, 0.33)	-0.03 (-0.13, 0.07)	-0.12 (-0.47, 0.23)	0.4875
	Follow-up	81	91	0.41 (0.33, 0.49)	0.43 (0.35, 0.51)	-0.02 (-0.13, 0.09)	-0.07 (-0.53, 0.39)	0.7429
Persistence Quality	Follow-up	78	99	3.42 (3.28, 3.55)	3.48 (3.36, 3.60)	-0.06 (-0.24, 0.13)	-0.10 (-0.42, 0.22)	0.5457
Parent report (CSBQ)	Post	92	103	14.90 (14.26, 15.54)	14.77 (14.12, 15.42)	0.13 (-0.79, 1.05)	0.04 (-0.27, 0.36)	0.7607
	Follow-up	86	100	15.56 (14.85, 16.27)	15.10 (14.40, 15.79)	0.46 (-0.54, 1.47)	0.19 (-0.22, 0.60)	0.3467
	Post	90	100	3.56 (3.13, 3.99)	3.72 (3.30, 4.14)	-0.16 (-0.77, 0.45)	-0.12 (-0.59, 0.34)	0.5991

Expressive (Emotion Understanding)	Follow-up	87	99	4.78 (4.43, 5.14)	5.04 (4.69, 5.39)	-0.26 (-0.76, 0.25)	-0.14 (-0.43, 0.14)	0.2925
Helpfulness (Empathy)	Post	83	98	6.75 (5.77, 7.72)	7.85 (6.92, 8.79)	-1.11 (-2.48, 0.26)	-0.30 (-0.66, 0.07)	0.1083
	Follow-up	77	97	8.40 (7.45, 9.35)	8.19 (7.30, 9.09)	0.20 (-1.11, 1.51)	0.05 (-0.37, 0.27)	0.7504
Altruism	Follow-up	83	97	3.57 (3.09, 4.06)	3.67 (3.23, 4.12)	-0.10 (-0.77, 0.57)	-0.05 (-0.37, 0.27)	0.7740
Theory of Mind	Post	90	100	0.41 (0.37, 0.44)	0.38 (0.35, 0.42)	0.02 (-0.03, 0.07)	0.14 (-0.15, 0.43)	0.3328
	Follow-up	85	99	0.49 (0.45, 0.53)	0.47 (0.43, 0.51)	0.02 (-0.03, 0.08)	0.13 (-0.22, 0.49)	0.4536
Prosocial (SDQ)	Post	91	103	7.60 (7.31, 7.89)	7.38 (7.11, 7.65)	0.22 (-0.19, 0.63)	0.15 (-0.13, 0.42)	0.2905
	Follow-up	86	100	7.83 (7.50, 8.17)	7.38 (7.07, 7.69)	0.45 (-0.01, 0.92)	0.35 (-0.01, 0.70)	0.0569
PARENT DISCIPLINE						**Adjusted difference (95% CI)**	**Standardised effect size (95% CI)**	
Harsh Psychological (parent report)(low good)	Post	91	103	1.72 (1.53, 1.92)	1.48 (1.30, 1.66)	0.24 (-0.03, 0.52)	0.27 (-0.03, 0.58)	0.0810
	Follow-up	87	100	1.50 (1.31, 1.70)	1.57 (1.39, 1.75)	-0.07 (-0.34, 0.21)	-0.07 (-0.38, 0.23)	0.6371
Non-Harsh-Negative (parent report)	Post	91	103	5.43 (5.11, 5.75)	5.19 (4.89, 5.49)	0.24 (-0.21, 0.69)	0.15 (-0.13, 0.43)	0.2911
	Follow-up	87	100	5.36 (5.01, 5.71)	5.06 (4.73, 5.38)	0.30 (-0.18, 0.79)	0.21 (-0.13, 0.55)	0.2209
	Post	91	103	4.81 (4.61, 5.00)	4.60 (4.41, 4.79)	0.21 (-0.07, 0.49)	0.31 (-0.10, 0.71)	0.1319

Non-Harsh – Positive (parent report)	Follow-up	87	100	4.88 (4.71, 5.05)	4.81 (4.65, 4.98)	0.07 (-0.17, 0.30)	0.08 (-0.19, 0.35)	0.5454
Physical Guidance ^	Post	91	102	0.50 (0.38, 0.63)	0.44 (0.31, 0.57)	0.06 (-0.12, 0.24)	0.15 (-0.32, 0.63)	0.4664
Verbal guidance	Post	91	102	1.60 (1.38, 1.82)	1.61 (1.39, 1.83)	-0.01 (-0.32, 0.30)	-0.01 (-0.44, 0.41)	0.9413

Adjusting for available baselines (and baseline by visit interaction where appropriate), child's age at each visit, child's gender, multilingualism and parental education. *adjusting for EYT baseline ^square root transformation used. $^{+}$ Follow-up visit only used in subjects where there was no video of book-sharing at post (3 participants). Intra-cluster correlation (ICC) was between 0 and 0.06 for all measures. For measures where a low value is a good, a negative standardised effect size is considered a positive effect of the intervention.

Appendix 4 Baseline scores for child measures and book-sharing interactions

	Intervention n	Intervention group mean (SD)	Control n	Control group mean (SD)
Expressive language (EYT)	109	13.93 (8.478)	108	18.94 (9.160)
Attention composite	81	-0.07 (0.522)	87	0.06 (0.473)
ECVT (%)	105	55.99 (23.671)	105	58.61 (23.419)
Play persistence	91	109.00 (35.981)	96	108.39 (31.592)
Play Quality	91	2.39 (0.821)	96	2.67 (0.681)
Task consistency (low good)	100	0.27 (0.097)	98	0.27 (0.107)
Parent report attention	110	3.45 (1.547)	108	3.36 (1.371)
Impulse control	104	0.00 (0.229)	101	0.05 (0.258)
Persistence Quality	104	3.04 (0.701)	106	3.07 (0.698)
Parent report (CSBQ)	110	13.64 (3.397)	108	14.01 (3.200)
Expressive (Emotion Understanding)	109	1.67 (2.023)	106	2.42 (2.133)
Helpfulness (Empathy)	104	6.67 (4.090)	105	6.73 (4.142)
Theory of Mind	109	0.23 (0.174)	106	0.33 (0.196)
Prosocial (SDQ)	110	6.84 (1.863)	108	7.26 (1.613)
Defiance (Don't touch)(low good)	110	0.81 (1.022)	108	0.71 (0.952)
Total difficulties (SDQ) (low good)	110	11.49 (4.887)	108	11.81 (4.833)
CSBQ (low good)	110	4.54 (2.577)	108	4.51 (2.290)
Sensitivity	110	2.73 (0.913)	108	2.85 (0.881)
Scale Reciprocity	110	2.69 (0.965)	108	2.79 (0.912)
Event Reciprocity	102	55.33 (23.018)	103	57.75 (20.068)
Cognitive scaffolding	98	-0.08 (0.682)	103	0.07 (0.751)
Mental state talk	107	2.70 (3.817)	108	3.75 (4.289)