Research Report: Outcomes of Talking Together

Evaluation and Results (oTTer): A randomised controlled feasibility trial

ISRCTN registration - https://doi.org/10.1186/ISRCTN13251954

Dr Dea Nielsen
University of York

Dr Katrina d'Apice
University of York

Dr Josie Dickerson
Bradford Institute of Health Research

Dr Maria Bryant
University of York

Rebecca Heald and Chloe Storr
BHT Early Education and Training

Dr Louise Tracey
University of York

Dr Claudine Bowyer-Crane
National Institute of Economic and Social Research

Table of Contents

Executive summary

Introduction

Concerns have been raised in recent years about the oral language development of young children, particularly in areas of disadvantage. Estimates suggest that up to 50% of young children from deprived backgrounds have speech, language and communication needs (SLCN). These concerns have led to the launch of campaigns such as Hungry Little Minds, and more recently some new guidance and an assessment tool published by Public Health England (PHE); recognition of the fact that early oral language development is a public health issue. What is less well understood is how best to support children who are identified as being at risk of language difficulties. However, one potential avenue of support is through improvement of the home learning environment.

The importance of the home learning environment was brought into sharp relief with the first national lockdown due to Covid-19 in March 2020. Schools and early years settings closed to all but children of key workers or vulnerable children, which meant the majority of children were unable to access the support and input they usually receive through attendance at settings outside the home. A second national lockdown began in January 2021. Once again, all schools closed, and families had to juggle work, childcare and home schooling. Early years settings remained open, but attendance rates did not return to pre-Covid levels with many families choosing to keep their young children at home. Clearly, therefore, support for families to provide a high quality home learning environment has never been more important.

Unfortunately, there are few rigorously evaluated programmes that focus on the home learning environment. The focus of this report is a home-based programme delivered by the charity BHT Early Education and Training. The Talking Together programme combines a universal language screening for all two-year-old children in the Better Start Bradford (BSB) reach area, with a targeted 6-week programme delivered to parents in the home. The programme consists of one-to-one visits over a six-week period delivered by specialist language development workers (LDWs). The programme is designed based on evidence of the importance of parent-child interaction, and aims to provide parents with the skills, knowledge and confidence to provide a language rich environment. Moreover, the programme has been adapted for delivery in an ethnically diverse community, with LDWs able to deliver the programme in many of the languages represented in the community, and interpreters employed to support families if necessary.

Talking Together is currently commissioned by BSB; a public health initiative funded by the National Lottery Community Fund. Together with Born in Bradford, BSB have established a new birth cohort; Born in Bradfords Better Start (BiBBS) enables evaluations of the impact of services on the outcomes for children and families.

Aims and Objectives

The primary research aim of the oTTer project was to establish the feasibility of a definitive randomised control trial (RCT) of Talking Together. There were two key objectives involved in meeting this aim; a) to assess the feasibility of conducting a trial to evaluate the effectiveness of Talking

Together including the acceptability and utility of the intervention outcome measures and b) to embed a qualitative evaluation within the oTTer trial to identify challenges with the implementation and delivery of the Talking Together programme as part of a trial. A secondary aim of the project was to understand factors in referral, uptake and attrition figures associated with Talking Together, which may inform the future provision of the service. The third aim of the project was to use the quantitative and qualitative data together to make an assessment of whether the Talking Together intervention shows evidence of promise.

Methods

For our feasibility study, we carried out a two arm randomised controlled feasibility study in a single site. 102 families were randomised in a 1:1 ratio to either an immediate intervention group, or a waiting control group.

We looked at routine monitoring data (i.e. number of referrals, rates of consent, completion rates, attrition rates, etc), as well as child language outcome data (potential primary outcomes for a definitive trials), and parent-level outcome data (potential secondary outcomes).

Potential Primary Outcome Measures

- Oxford CDI- Short: This is a vocabulary checklist used to assess children's lexical knowledge.
- WellComm Early Years: This tool uses a combination of observation, direct assessment, and parent report to gain a holistic picture of children's language skills.

Potential Secondary outcome measures

- Maternal Object Relations Scale (MORS): a measure of attachment and parent/carer and child relationship
- Home learning environment questionnaire (HLEQ): This measure assesses the frequency with which families engage in eight types of language enriching activities in their home.
- Strengths and Difficulties Questionnaire (SDQ): The SDQ is a well-validated and widely used measure of children's emotional and behavioural wellbeing.

Additionally, qualitative data was collected through interviews with practitioners and families to explore barriers and facilitators to participating in the trial.

We used a RAG (red, amber, green) rating for progression criteria that focused on recruitment, adherence and attrition.

To explore factors in referral, uptake and attrition, data from the BiBBS cohort was linked to data from families who attended the two-year universal language screening. A series of regression analyses were conducted with ethnicity, first language and level of education as predictor variables.

Key Findings

Progression Criteria

- Recruitment: 74% of families referred into Talking Together were eligible for the trial, and 62% gave consent to take part, putting recruitment in green in terms of progression criteria.
- Adherence to the assessment timeline set out in the protocol ranged from 51% (red) to 89% (green).
- We had a 32% attrition rate, which we classed as red according to our progression criteria.

Identification of Primary and Secondary Outcome Measures

- CDI Understanding was identified as our primary outcome measure. Data showed that this measure had a high level of completeness, was reliable and responsive.
- MORS was identified as our secondary outcome measure. The MORS measure showed a high level of completeness across all time points. The MORS Warmth subscale was the most responsive to change over time of all the non-language measures and correlated well with the language measures. Reliability of the MORS Warmth subscale was less than ideal. However, on balance this measure appeared to be the most appropriate for capturing the impact of the programme on parents.

Barriers and Facilitators

- Qualitative interviews suggested that the programme was positively received by families. There was some concern from the LDWs over the use of a waiting control group, which resulted in some eligible families not being offered the trial. Some LDWs were also concerned about the increased administration involved in being part of the trial. Training and teamwork were key facilitators in overcoming some of these concerns.

Predictors of referral, uptake and attrition

- There were no significant predictors of referral or attrition. This suggests maternal education or speaking English as a First Language did not impact on how families were referred into the intervention or how likely they were to complete the intervention.
- Regression analyses revealed that uptake, or acceptance of the Talking Together programme, was predicted by the variable First Language. Families who had English as a First Language were more likely to accept the offer of Talking Together than families for whom English was not their First Language. This is not surprising, given that the intervention is primarily run in English and families with English as a First Language are likely to find it easier to access this service than families with limited or no English language abilities. However, it does suggest that there may be parts of the community that are underserved due to their language background.

Evidence of promise

- Based on all the available evidence, including group differences in favour of the intervention group on the identified primary and secondary outcome measures, the acceptability of the intervention, as well as the experience of participants that the intervention made a difference to parents and children, it was considered that this intervention shows evidence of promise.

Conclusion

Referral rates into Talking Together indicate that it is a much-needed service, and the results of this study indicate that is positively received by the community it serves. Interpretation of our results against progression criteria would suggest that full trial of the programme is feasible with some adaptations particularly relating to reducing attrition.

Next Steps

The results of the study provide clear support for further research into the effectiveness of Talking Together. This feasibility and pilot study is one of the first steps on the path to an effectiveness evaluation. The next step should be an efficacy evaluation, using an RCT or quasi-experimental design.

In addition, we need to explore the replicability of the programme beyond the specific areas of Bradford where it has been running to date.

Acknowledgements

We would like to thank all the families and Language Development Workers who took part in this research. We would also like to thank Amirah Akthar, Lauren Burke, Sameera Ali, Nimarta Dharni and Rukshana Rasheed for their support. Finally, we would like to thank the Nuffield Foundation for funding the research. The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. The Nuffield Foundation has funded this project, but the views expressed are those of the authors and not necessarily those of the Foundation. More information is available at www.nuffieldfoundation.org.

Introduction

Oral Language and the Home Learning Environment

There has been increasing momentum in the promotion of oral language as a fundamental foundation for children's educational achievement, social and emotional development, and even later employment opportunities. Reports from Bercow[1], The Communication Trust[2], and the Early Intervention Foundation[3] for example have shown that a large number of children in the UK are starting school with language skills below that expected for their age. Speech, language and communication needs (SLCN) are particularly prevalent in areas of disadvantage; it is estimated that language difficulties affect over 20% of children from more deprived backgrounds[4]. These concerns have led to the establishment of campaigns such as the Government's Hungry Little Minds campaign. More recently Public Health England have published a set of guidelines for professionals working with very young children that detail how best to support language and communication skills[5]. A new assessment has been published alongside these guidelines for use by health visitors and other professionals to identify children in need of support as early as possible[6]. While this is certainly a positive move, the issue of how best to support those children at risk, and indeed who should provide that support is still unclear. What is clear however, is that the home learning environment (HLE) is the bedrock of children's early language development[7,8]. Therefore, efforts to support children at risk of

language difficulties could be effectively targeted at working with families who need help in developing the knowledge and skills necessary to support their children's language development.

The importance of a high-quality home learning environment was brought into sharp relief with the closure of schools and early years settings in March 2020 as a result of the COVID-19 pandemic. Families were being asked to educate their children at home whilst juggling work commitments, and for many families, health issues and financial concerns[9]. Unfortunately, reports have emerged suggesting that this situation has led to increasing inequalities[10]. Switching to remote learning left many children disadvantaged by the "digital divide", with children from more affluent backgrounds having better access to laptops, tablets and smartphones than those from less affluent backgrounds[11]. In addition, children who accessed additional support through attending settings were no longer in receipt of that support, and children with difficulties emerging during lockdown may have gone unnoticed because contact with services was limited. For example, a report from the children's commissioner highlighted that face to face visits with health visitors and social workers were to be carried out virtually, if at all, making it extremely difficult to identify children with additional needs, particularly in the area of language and communication[12]. Moreover, the pandemic saw a rise in parental mental health issues, domestic abuse and children living in poverty; all factors that have a significant impact on a child's development[12]. As settings reopened, it became clear that children's language and communication skills suffered as a result of the lockdown. Recent Ofsted briefings have shown that children returned to reception classes with language and communication skills below that expected for their age[13]. In addition, a survey of early years providers suggested that language and communication skills were a priority moving forward, second only to personal, social and emotional development[14]. The nation embarked on a second lockdown in January 2021. Once again, all schools closed, meaning that families were once again required to juggle, work, childcare and home schooling, increasing stress in the home. Moreover, while early years settings remained open, attendance rates did not return to pre-Covid levels[15]. Clearly, therefore, support for families to provide a high quality home learning environment has never been more important.

Unfortunately, there are few rigorously evaluated interventions, using home-visiting or other designs, for young children identified as being at risk of language delay[16]. Campaigns such as Hungry Little Minds and Tiny Happy People provide resources for families that aim to increase their knowledge of language development, and give them ideas for activities that will improve their interactions with their babies and young children. But without direct support, the messages contained in these resources may not be implemented effectively. This report details a feasibility evaluation of the Talking Together programme; a home-based language and communication support programme developed and run by BHT Early Education and Training (BHT) in Bradford, which seeks to fill this gap. The team at BHT worked directly with parents to support the home learning environment of children from disadvantaged backgrounds, and culturally and linguistically diverse families. The Talking Together programme combined a universal screening assessment for all families with children aged two years of age, with a programme of one-to-one visits in the home over a six week period for families identified as in need of support. The programme itself was delivered by specialist language development workers (LDWs) who had extensive training in children's language development. It was designed based on evidence of the importance of parent-child interaction, and aimed to provide parents with the skills, knowledge and confidence to foster a language rich environment. Although when the intervention was developed it mainly served a white English-speaking British community living in relatively high levels

of deprivation, it was successfully adapted to serve the ethnically and linguistically diverse BSB community. LDWs were able to deliver the programme in many of the languages represented in the community, and interpreters were employed to support families if necessary.

Better Start Bradford

Talking Together was commissioned by Better Start Bradford (BSB); a public health initiative funded by the National Lottery Community Fund which focused on providing preventative early years interventions to an ethnically diverse community living in inner-city area of Bradford. Launched in 2015, BSB commissioned numerous programmes that focussed on three key areas; social and emotional development, nutrition and obesity, and language and communication. The programmes were aimed at pregnant women, and families with young children. A wide range of services were commissioned including antenatal and midwifery, breastfeeding support, parenting programmes, healthy eating programmes, and programmes encouraging the use of outdoor spaces. To support language and communication, BSB commissioned a book gifting scheme and a programme of CPD for early years practitioners as well as the Talking Together programme.

BSB partnered with the long-running Born in Bradford Birth Cohort study to establish a new birth cohort; Born in Bradford's Better Start (BiBBS)[17]. Families within BSB were invited to take part in BiBBS at routine pregnancy appointments. The aim of BiBBS was to follow the journey of families using routinely collected data from health, education and BSB projects to enable evaluations of BSB projects on outcomes for children and families.

Aims and Objectives

The primary research aim of the oTTer project was to establish the feasibility of a definitive RCT trial of Talking Together. There were two key objectives involved in meeting this aim; a) to assess the feasibility of conducting a trial to evaluate the effectiveness of Talking Together including the acceptability of the intervention outcome measures, and whether they suggest the intervention shows evidence of promise in terms of group differences and b) to embed a qualitative evaluation within the oTTer trial to identify challenges with the implementation and delivery of the Talking Together programme as part of a trial.

The research questions for the feasibility study (objective a) were:
1. What are the recruitment and retention rates of Talking Together established by the number of participants who were identified, eligible, approached, consented, randomised, completed the programme, and followed up six months after baseline?
2. How representative are the trial participants compared to the wider population receiving the intervention, based on key demographic indicators?
3. What are the most appropriate outcome measures for a future definitive RCT, considering group differences, the acceptability, reliability, data quality (completeness), and responsiveness of administered measures?
4. What is the sample size needed for a definitive trial based on intervention completion and attrition rates, along with outcome data, group differences and variability between study arms?

5. Was the intervention delivered with fidelity to the standardised procedures as measured by assessing the intervention content, and the frequency and duration of support received by participants?
6. What are the time and resources required to train practitioners to administer the intervention, and how do these relate to resource requirements for definitive RCT development?

The research questions for the qualitative evaluation (objective b) were:
1. How acceptable are the intervention and trial procedures for practitioners and families, including randomisation and completion of outcome measures?
2. What were the barriers and facilitators to the intervention and the trial?

A secondary aim of the project was to understand factors in referral, uptake and attrition figures associated with Talking Together, which may inform the future provision of the service. For example, if particular families are at risk of refusing or not completing the programme, BHT could develop strategies to mitigate this outcome. Our objective was to link the data from the BiBBS cohort to families screened for referral into Talking Together and to explore the links between level of education, ethnicity and language, maternal attachment and socioeconomic status on referral, uptake and attrition figures. The research question for this objective was:

1. How does level of education, ethnicity, first language, maternal mental health and SES relate to referral, uptake and attrition for the Talking Together programme?

The third aim of the project was to use the quantitative and qualitative data together to make an assessment of whether the Talking Together intervention shows evidence of promise. The research question for this objective was:

1. Does the evidence collected from quantitative and qualitative data show evidence of promise that the Talking Together intervention has a positive impact on parent and child outcomes?

Methods

Feasibility Study

Study design

The main component of this work was a two-armed individual randomised feasibility study in a single site. Participants were randomised in a 1:1 ratio to either an immediate intervention group, or a waiting control group. Participants in the immediate intervention group began the intervention shortly after allocation, while participants in the waiting control group began the intervention (if still appropriate) approximately six months after randomisation. The study protocol has been published, and is available here[18] https://doi.org/10.1186/s40814-019-0498-2.

Implementation and feasibility outcomes were assessed using routine monitoring data (i.e. number of referrals, rates of consent, completion rates, attrition rates, etc.), as well as child language outcome data (potential primary outcomes for a definitive trial), and parent-level outcome data (potential

secondary outcomes). Additionally, qualitative data was collected through interviews with practitioners and families participating in the trial.

Setting and participants

BSB serves three areas of the city. These areas were specifically selected for this programme due to their high levels of deprivation, and according to IMD figures they represent some of the most deprived areas of England. Within the BSB community, approximately 37% of children did not achieve a Good Level of Development in the EYFSP in language and communication in the 2018/19 academic year, suggesting that there is a demonstrable need for targeted support for children's language skills in the early years[19].

It was estimated that during the recruitment phase, approximately 670 families would be seen for the Universal Language Assessment (screening), and of these, approximately 250 families would be referred on to the Talking Together programme. All families who were offered the intervention during the recruitment phase were assessed for eligibility to participate in the feasibility study.

Interventions

Talking Together intervention

The Talking Together programme comprised two components; the Universal Language Assessment and the targeted Talking Together intervention. Language Development Workers (LDWs; early years practitioners with extensive training in children's early language and communication development) delivered both components of the programme, and the whole service was funded by BSB.

The Universal Language Assessment was offered to all families of a child aged two within the BSB reach area. Data for those families with consent to data sharing with BSB was passed onto the Talking Together service provider, BHT Early Education and Training (BHT) from Bradford District Care Trust. BHT then invited families to receive their screening via a letter, and LDWs scheduled an appointment to visit the family in their home. During this visit they administered both the Universal Language Screener and the Oxford CDI[20] (see Outcomes section), and observed the child, the parent, and the home environment. On the basis of these measures and their observations, LDWs made a decision about whether there was a concern about the child's language and communication development, and referred into the intervention accordingly.

The Talking Together intervention aims to provide parents with a foundational understanding of children's language development, and to improve parent-child interaction and the home learning environment. These improvements are then believed to lead to advances in children's language skills (see Theory of Change, in Figure 1).

The programme consisted of six weekly sessions, delivered one-to-one with parents and children in the home by LDWs. The sessions focus on five topic areas, including *what is communication*, *the importance of play*, *attention and listening*, *turn-taking*, and *praise and encouragement*, each of which is covered in one session. The final session is used to summarise the content covered, and can be adapted to the needs and interests of individual families. Sessions begin with LDWs providing pre-specified information on the week's topic, and then LDWs work with parents to consider how to support their child with skills related to the week's topic in a way that works for the family's environment and interests. At the heart of the programme is the importance of taking an assets-based approach, to ensure that parents feel supported rather than judged and to build on the skills and resources that families already have available to them. As such, while the programme is manualised and fidelity to the delivery of the core content is monitored, LDWs are encouraged to personalise the content for each family to make it as relevant and accessible as possible. LDWs provide personalised ideas for activities, bring along resources for the family to enjoy and be inspired by, and leave activities that parents complete between sessions. The sessions also allow for LDWs to model good adult-child interaction while observing parent's behaviours, ensuring they praise their strengths while gently suggesting alternatives for less optimal behaviours. Parents are encouraged to discuss their own concerns so that LDWs can work with them to troubleshoot any challenges they face (for example, children's behaviour or discipline). This approach also ensures that the programme is language general, and is appropriate for use with any family, regardless of their home language.

Assessments of both parent and child outcomes were carried out in the first (T2) and last session (T3) of the programme by LDWs as an integrated part of these sessions (see Outcomes). At the follow up assessment point (T4), the assessments were carried out by research assistants (RA), who attended with an LDW to provide continuity for families.

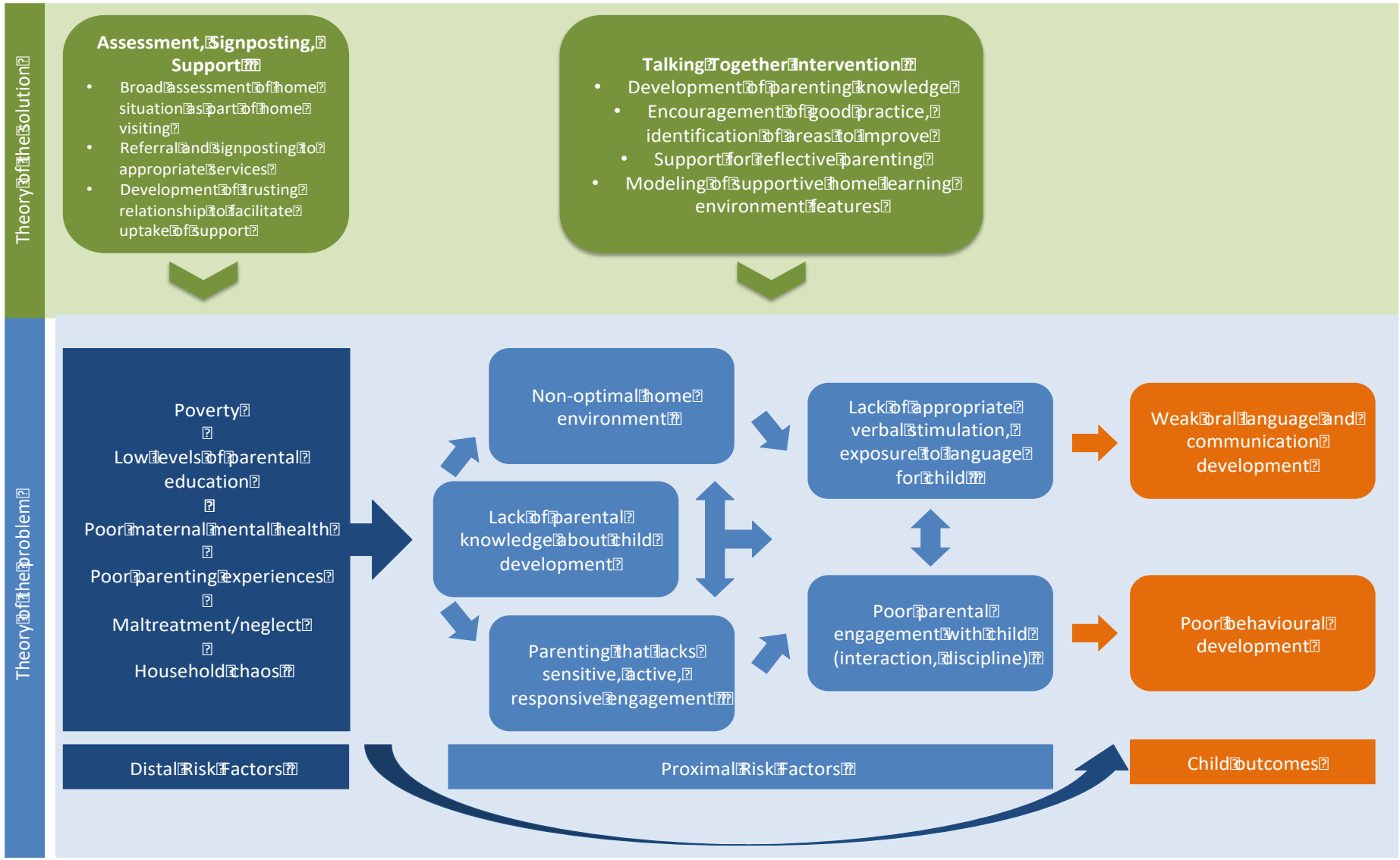

Figure 1. Talking Together theory of change.

Waiting control

During the first assessment session (T2), families in the waiting control group received a packet of information, resources, and activities that they could use with their child. The RA explained the contents to the family, and ensured they understood all the contents and where they could look for additional information independently. An RA then saw families again at the second assessment session (T3), and finally at the follow up assessment (T4) the measures were completed by an LDW in the presence of an RA. This allowed the LDWs to assess whether the intervention was still needed for child. As such, although it was not possible for the assessors to be blinded, for both arms of the trial T4 assessments were carried out by someone who had not been involved with the child's treatment previously.

Within the community at the time of the trial, the provision for children's language and communication development was limited. Health visitors with concerns about children's language development at the two-year assessment could refer to Speech and Language Therapy Services, as could children's early education and childcare settings. Beyond that, another BSB project ran story and rhyme sessions open to all families, and families were encouraged to download the Bradford 50 Things to do before 5 app. No other dedicated language support was available.

Sample Size

As this was a feasibility study and not an effectiveness evaluation, the study was not powered to detect significant effects. The sample size was based on previous guidance for feasibility studies, which suggests 30 participants per arm [21,22]. We aimed to recruit 120 participants at screening (60 participants per arm) to allow for attrition.

Eligibility criteria

Families were invited to participate in the study if they met the following inclusion criteria:

- they had been referred to Talking Together by a LDW following the Universal Language Screening
- they consented to receive the intervention in their home
- they lived within the reach area of BSB
- their target child was aged between 2 to 2.5 years
- they nominated a specific family member to receive the entirety of the intervention
- they consented to randomisation, and accepted that if they were allocated to the control group they would wait six months to receive the intervention and would be visited for additional data collection during the wait
- they spoke primarily English, Urdu, or Punjabi with the target child

Families were excluded from the study if they met any of the following criteria:

- their target child had a known sensory impairment or developmental disorder
- their referral into Talking Together came from a source other than an LDW (i.e. safeguarding authorities)

- they were unable to confirm a specific family member to participate in the entirety of the intervention

Randomisation, Allocation, and Blinding

Referrals from families with informed consent to participate in the trial were randomised (1:1) to either immediate intervention or waiting control. Minimisation was used to ensure that the two groups were balanced on (a) language of delivery (English or not English), (b) Children's Centre reach area (Area A or Area B), and (c) whether multiple children would be present during the delivery of the intervention (yes or no). Children's Centre reach area was included based on previous data suggesting that need for and uptake of the intervention was different in this area compared to others within the reach area. The randomisation process was carried out by a statistician within the data management (DM) team at BiB. Eligible referrals were entered onto the SystmOne database, which the DM team then accessed once a week in order to randomise. Each referral allocation was then recorded on SystmOne (a centralised Electronic Health Record system used to allow continuity and communication between multiple health and care providers), and accessed by the senior BHT staff. Families were told of their treatment allocation by LDWs (invention group) or RAs (waiting control) when contacted to arrange their first session.

It was not possible to blind LDWs and RAs to treatment allocation due to their need to deliver the intervention and assessment sessions. The research team were also aware of treatment group for anonymous participants as this was necessary for monitoring the trial and the data analyses, and one specific member of the research team was unblinded to treatment conditions in order to conduct participant interviews.

Objectives

The primary research aim of this study was to determine the feasibility of conducting a definitive RCT of the Talking Together intervention. To this end, the primary feasibility objective was recruitment rate and retention in the trial. This was assessed through counts of the number of families that:

- received the universal screening
- were eligible to receive Talking Together
- were offered Talking Together
- accepted Talking Together
- were eligible for the feasibility study
- were not eligible for the feasibility study and the reason why
- were approached to take part in the feasibility study
- consented to take part in the feasibility study
- The timing, number, proportion, and reason for withdrawals from the intervention
- The timing, number, proportion, and reason for withdrawals from the trial

In addition to the primary objective, there were a number of secondary objectives that were important to understanding future trial feasibility. These were:

- To understand the representativeness of the trial participants compared to the wider population of children receiving the intervention, compared on:
    - Age
    - Gender
    - Ethnicity
    - Home language
- To identify the most appropriate outcome measure for a future trial by considering data collection and completeness, including:
    - Data completeness for all participants at each time point (e.g. the number of participants with quantitative data)
    - Data completeness for all outcome measures at each time point (e.g. proportion of missing item-level data)
    - Internal reliability of the outcome measures
    - Standardised Response Means for the outcome measures
    - Differences between groups at the beginning and end of the intervention
    - Correlations between the outcome measures
- To estimate the sample size needed for a future definitive trial
    - The difference between the immediate intervention and waiting control group (means, 95% confidence intervals) at the 6-month follow up on the Oxford CDI[20] and the WellComm[23]
- To understand if the intervention and trial were delivered with fidelity to standardised procedures
    - Number and timings of session delivery for all participants
    - Number and proportion of participants who complete the interventions
- To understand the resource requirements for the current trial and how these relate to the requirements of a definitive trial using:
    - Estimates of the time and resources necessary to train LDWs to deliver Talking Together
    - Counts of the number of fully trained LDWs
    - Quality assurance data in the form of staff observations of adherence to standard intervention content
- To understand the acceptability and barriers and facilitators to the intervention and the trial using:
    - Qualitative interviews with parents participating in the trial
    - Qualitative interviews with practitioners delivering the intervention and trial

Assessment and Outcome measures

Screening measures

The Universal Language Assessment is carried out 1:1 in families' homes by a LDW. There are two elements to this assessment.

- BHT Language Screener: This tool was originally developed by BHT in collaboration with a local Speech and Language Therapy service, and later updated with input from academic partners.

LDWs present parents/carers with ten statements about early communication and language skills, and parents answer whether their child is showing this skill often (two points), sometimes (one point), or not yet (no points). Parents are asked to consider all of their child's languages, so the assessment is suitable for both monolingual and bilingual children learning both English and non-English languages. As LDWs are assessing for multiple risk factors for early language weakness through both these assessments and their observations, the screener also asks LDWs to record their main reason for referral to the intervention (language and communication, child behaviour, parent behaviour, home learning environment, or supporting a family with complex needs).

- Oxford Communication Development Inventory-Short (CDI-Short)[20]: This is a vocabulary checklist used to assess children's lexical knowledge. LDWs read out a list of 100 words, and parents are asked for each word whether their child can understand it and whether they use it productively (e.g. speak it).

Intervention outcome measures

Children in both arms of the study were assessed on intervention outcome measures three times; pre-test, post-test, and follow up to determine the feasibility of these measures for use in a future definitive trial (see Figure 2). Assessments were administered by an LDW or RA in the family's home. These assessments were selected based on considerations of level of expertise required to administer the assessment, the ease of administration, the time to administer, the cost, and the appropriateness of the measure for monolingual and bilingual children speaking different languages.

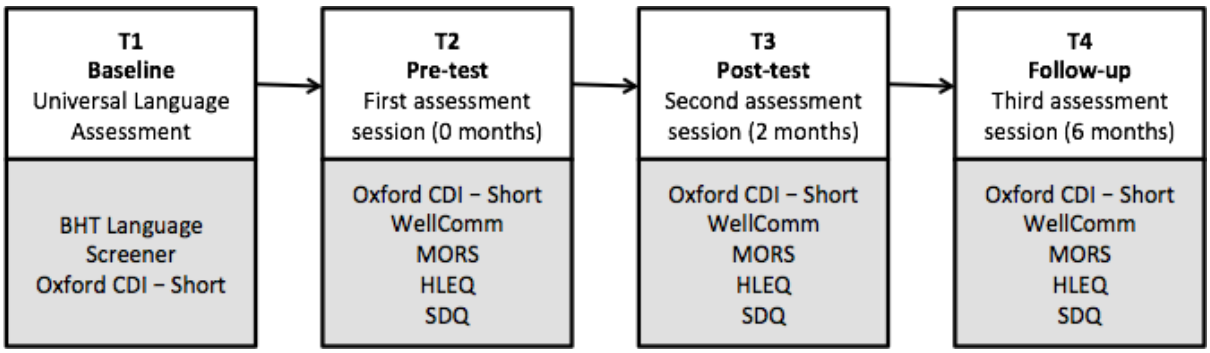

Figure 2. Participant timeline and schedule of assessments (white box), and assessments administered at each timepoint (grey box).

Potential primary outcome measures - child outcomes

- Oxford CDI- Short[20]: This assessment was administered in the same way as during screening (see above).
- WellComm Early Years[23]: This tool was designed to be used by anyone working with early years children and is both relatively short and simple to administer. It uses a combination of observation, direct assessment, and parent report to gain a holistic picture of children's language skills. The assessor starts with the 10-question section of the measure corresponding to the child's age (sections cover 6-month age bands), and each score receives a RAG rating of

green, amber, or red. Assessors then work backwards through the age bands until the child's score is categorised as green.

Potential secondary outcome measures - parent outcomes

- Maternal Object Relations Scale (MORS)[24]: This 14-item measure of attachment and parent/carer and child relationship asks parent/carers to rate on a six point scale (0-5) statements about their child. Assessments were completed with the support of an LDW, who read out the questions for the parent/carer to answer. This tool has previously been validated for use with similar populations. Scores are out of a possible 70 points.
- Home learning environment questionnaire (HLEQ)[25]: This measure assesses the frequency with which families engage in eight types of language enriching activities in their home. The assessment was completed with the support of an LDW, and scores are out of a possible 56 points.
- Strengths and Difficulties Questionnaire (SDQ)[26]: The SDQ is a well validated and widely used measure of children's emotional and behavioural wellbeing. For the purposes of this study, only two of the five subscales (hyperactivity and conduct) were chosen for administration. This was a pragmatic choice to decrease the length of the testing battery, and because these subscales were considered the most relevant to assessing the needs of the children receiving the intervention. This short version of the measure contained ten questions, and parents/carers completed the measure by answering the questions read out by an LDW.

Process evaluation

The research team conducted 60-minute interviews with practitioners and families. Twelve practitioner interviews and 23 parent interviews were conducted, audio recorded, and transcribed. A further 2 parent interviews were conducted but not audio recorded, in one of these cases the interviewer's notes were used for subsequent analysis. The transcriptions and interviewer's notes were analysed using NVivo 12 Pro. Practitioner interviews were conducted at BHT's offices, while parent interviews were conducted in families' homes.

Coding

For the parent interviews, a coding framework was developed based on themes that emerged during the interviews. The codes were then refined by two researchers to ensure that the codes aligned with the research questions.

For the staff interviews, codes were identified by one researcher based on the research questions that we wanted to answer. To ensure comprehensive coverage of the material, the theoretical domains that were used to create the interview questions were included as separate codes in the framework. A different researcher then reviewed the coding framework and assessed the codes over three transcripts to ensure they were suitable. Any comments or issues were discussed between the researchers, and the final coding framework was agreed upon. One of the researchers then coded the remaining transcripts.

Analysis

A deductive approach was applied to the analysis of the transcripts, it was pre-specified almost coding to the research question. We then extracted the information that related to each code and examined the patterns in the participants' responses.

Progression Criteria

Our progression criteria are based on recruitment rates, protocol adherence and attrition.

Recruitment - our estimates were that, of families offered Talking Together, at least 60% would be eligible for the trial, with an anticipated consent rate of 50%. To meet green criteria, at least 60% of parents offered Talking Together would be eligible to take part, 50-60% would be classified as 'amber' and a recruitment rate of below 50% would be classified as red. For consent, 50% and above will be considered green, 40-50% amber and below 40%, red.

Protocol adherence - assessments of intervention group families and waiting control group families should take place at specific time points to ensure they align. To run a trial using the current design, T2 assessments should be run within one month of the Universal Language Screener (T1); T3 assessments should be run 6-10 weeks following T2, and T4 assessments should be run 5.5 to 6.5 months following T3. Progression decisions were therefore based on percentage adherence rates, i.e. 80% = green, 60–80% = amber, and less than 60% = red.

Attrition rates - predicted attrition rates were based on previous attrition rates from the Talking Together programme. Progression decisions were designed to be based on the proportion of the recruited sample attended the 6-month follow-up (T4), 80% = green, 70% = amber, and below 70% = red.

BIBBs Cohort

Study Design

The aim of this part of the project was to explore potential factors related to referral, uptake and attrition for the Talking Together programme. An exploratory study was carried out linking data from families who had been screened using the universal language screener, with data from the BiBBS cohort[17]. A series of regression analyses were conducted to answer the research question "How does level of education, ethnicity, first language, maternal attachment and SES relate to referral, uptake and attrition for the Talking Together programme?"

Sampling

Recruitment to the BiBBS cohort started in January 2016 and is ongoing. All pregnant women due to give birth at Bradford Teaching Hospitals NHS Foundation Trust (BTHFT) who lived in the BSB reach area were eligible to take part, and their partners were also invited to participate. The exception to

this was any women who planned to move away from Bradford prior to giving birth, and women who did not want the research team to approach their partner. Recruitment took place primarily through attendance at the Glucose Tolerance Test (GTT), or alternatively through other routine pregnancy appointments. At recruitment, trained researchers carried out a baseline questionnaire with mothers that included questions on households, SES, neighbourhood, language and communication, ethnicity, demographic data and health and wellbeing, and mothers also provided a blood sample, urine sample and anthropometric measurements. To date, over 3000 mothers have been recruited, along with more than 300 partners, and over 2,700 babies have been born to the BiBBS cohort.

Data linkage

Data from families screened for Talking Together during the oTTer recruitment period were linked to the BiBBS database by the data team at Better Start Bradford Innovation Hub (BSBIH). Data from the BiBBS database for the variables level of education, ethnicity, first language, maternal attachment and SES was then obtained. Unfortunately, the level of missing data for the maternal attachment and SES variables meant it was not possible to include them in the analyses.

Ethnicity was recorded in line with BiB policy combining values for any cell with small numbers (<5) to protect anonymity. Ethnicity was coded into four categories: White British, White Other, South Asian, and Other.

First language was a binomial variable because participants answered either *yes* or *no* to the question "Is English your first language?"

Level of education was a derived categorical variable equating educational qualifications from different countries.

Results

Changes to protocol

The processes and procedures set out in the protocol were overwhelmingly adhered to, meaning it was possible to collect the vast majority of the data as expected. The exception to this was the aim to collect data on LDWs delivery of the intervention sessions in order to document fidelity to standardised intervention procedures. The service provider did collect this data, but it was not saved appropriately, and so could not be collected by the research team.

There was also a small change to the randomisation procedure within the first two weeks of the trial. The original randomisation process did not protect against multiple allocations of the same participants to conditions, resulting in one participant being allocated twice, once to each group. The participant's first allocation was honoured, and the randomisation programme was amended to ensure this did not happen again.

Finally, the original progression criteria were slightly amended for clarity. A record of these changes can be found of the ISRCTN trial registration (http://www.isrctn.com/ISRCTN13251954).

Feasibility Study

1. What are the recruitment and retention rates of Talking Together (for the oTTer trial) established by the number of participants who were identified, eligible, approached, consented, completed the programme, and followed up six months after baseline?

The target recruitment figure for the trial was 120 participants, with a minimum sample size set at 60. Total recruitment into the study was 102 participants. Figure 3 shows the cumulative recruitment figures by month between 10th October, 2018 and 14th June, 2019.

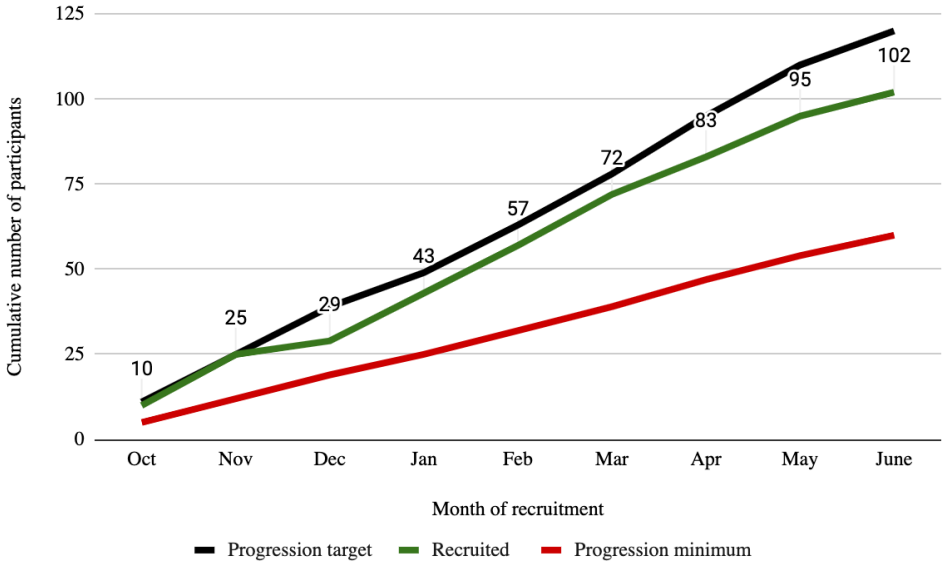

Figure 3. Cumulative recruitment rate

The black line represents the monthly recruitment figure that would be needed to reach the target sample size set out in the progression criteria (*n* = 120), while the red line represents the monthly figure that would be necessary to achieve the minimum 60 participant sample size. The green line represents the actual recruitment rate, and shows steady gains in participant numbers throughout the 9-month period with no notable fluctuations. On average, 11 participants consented to participate each month. Figure 4 shows the CONSORT diagram of participant flow through the trial from screening through data analysis. During the recruitment period, 608 families were screened for referral into Talking Together. Of these, 336 were offered the intervention (55% referral rate), and 264 of those families accepted the offer (79% acceptance rate). Of these families, 222 were assessed for eligibility for the trial, while 42 had missing eligibility data. Of the 222 families assessed for eligibility, 58 were not eligible for the trial, and 62 did not consent to being in the trial.

The eligibility rate was calculated by dividing the number of families eligible to take part in the trial (164) by the number who were assessed for eligibility to the trial (222), excluding missing data. This

resulted in an eligibility rate of 74%. Similarly, the consent rate was calculated by dividing the number of families who consented to be involved (102) by the number of families eligible to take part (164). As such, the trial had a consent rate of 62%.

In total, 102 families consented to take part in the trial. Fifty-two were randomly allocated to the intervention condition immediately, and 50 were allocated to the waiting control group. In the intervention group, 41 families (79%) went on to receive the first session of the intervention. Between randomisation and intervention start, six families withdrew from the trial, one was identified as an eligibility violation, one was considered to no longer need the intervention (and were therefore not eligible for the trial), and three became uncontactable. Thirty-six families finished the intervention, while five families did not. Of these, one withdrew from the trial, one moved out of the area, and three became uncontactable. Finally, 33 families (63%) were seen at the final assessment point, while three could not be followed up. Of these three, one was discovered to be ineligible, one moved out of the area, and one became uncontactable.

In the waiting control group, 45 (90%) of the allocated 50 families were seen for the first assessment session. Four families withdrew before the first session, and one was discovered to be an eligibility violation. Thirty-eight families were then seen at the second assessment point, while five withdrew before this point, one was discovered to be an eligibility violation, and one became uncontactable. Thirty-six families (72%) were seen for the final assessment point, while one family withdrew before this point, and another one family became uncontactable.

All families with available data were included in the final analysis (immediate intervention n = 33, waiting control n = 36).

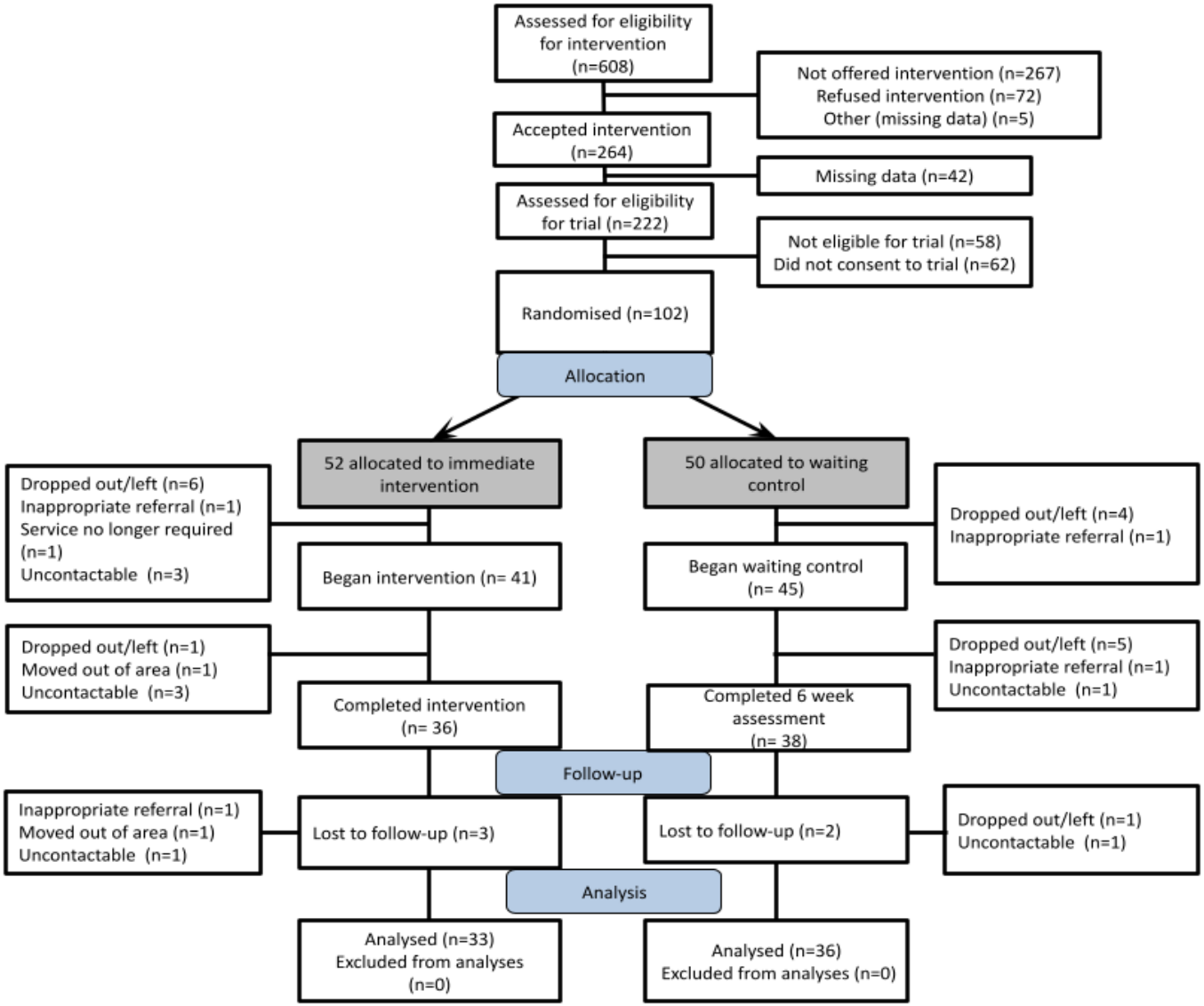

Figure 4. CONSORT diagram showing participant flow through the project

2. How representative are the trial participants compared to the wider population receiving the intervention, based on key demographic indicators?

Given that the eligibility criteria for the trial were more stringent than for the intervention in general, it was important to consider how representative participants of the trial were of the overall population of children receiving TT, as well as ensuring demographic similarity between the two arms of the study. The demographics for trial and intervention participants are shown in Table 1. Children taking part in the trial were compared to all other children receiving TT during the recruitment period (10/10/2018 - 15/6/2019) on their age, gender, ethnicity, and home language. Trial participants were on average 25.83 months old at T1, which was similar to other TT participants (26.51 months). There was a larger proportion of female children in the trial sample (47%) as compared to the other TT participants (42%), although this was due to a larger proportion of females in the waiting control arm (54%) as compared to the intervention arm (40%) of the trial. There was also a larger proportion of children from Asian/Asian British backgrounds in the trial (74.51%) as compared to other TT participants (66.05%), and also a larger proportion of White British families in the trial (13.73%) than in TT in general (4.94%). Conversely, all other ethnicities were underrepresented in the trial (11.76%) as compared to other TT participants (29.01%). Aside from the larger proportion of female participants in the waiting control group, the participants in the two arms of the study were very similar on these demographic variables.

Table 1. Characteristics of the whole oTTer sample, the Intervention and Waiting Control Groups, and all other families referred into Talking Together

		oTTer sample	Intervention	Waiting control	Other TT recipients*
N		102	52	50	162
Mean child age in months (SD)		25.83 (0.75)	25.87 (0.82)	25.8 (0.67)	26.51 (2.73)
Child gender (% female)		47	40	54	42
Child ethnicity (%)	Asian/Asian British	76 (74.51)	38 (73.08)	38 (76.00)	107 (66.05)
	White British	14 (13.73)	9 (17.31)	5 (10.00)	8 (4.94)
	Other	12 (11.76)	5 (9.62)	7 (14.00)	47 (29.01)
Home language	English	59 (59.00)	31 (59.62)	28 (58.33)	52 (33.12)
(%)	Urdu	16 (16.00)	8 (15.38)	8 (16.67)	21 (13.38)
	Punjabi	14 (14.00)	9 (17.31)	5 (10.42)	16 (10.19)
	Other	11 (11.00)	**	**	68 (43.31)
	Missing data	2	0	2	5

Note. *participants who received Talking Together who were ineligible for or did not consent to the trial. **Any data where n < 5 was withheld or collapsed into the group 'other' to protect anonymity.

With regards to home language, the proportion of children with English as their home language was notably higher in the trial (59%) compared to the other TT participants (33.12%). However, the proportion of Urdu and Punjabi speaking children was similar across the two groups, at 16% and 14%

respectively in the trial participants, and 13.38% and 10.19% in the other TT participants. All other languages, including Bengali, Polish and Hindko, were spoken by a lower percentage of trial participants compared with other TT participants, at 11% and 43.32% respectively.

3. What are the most appropriate outcome measures for a future definitive RCT, considering group differences, the data quality (completeness), reliability, acceptability and responsiveness of the measures?

There were a number of important considerations when identifying the most appropriate outcome measure for future evaluation of the intervention. Specifically, data completeness, measure reliability, the relationship between language and non-language measures, and group differences on the measures were all assessed and used to select the most promising primary and secondary outcome measures.

Data completeness

Table 2. Data completeness in total number of completed assessments at each time point for each group (percentage completion rate in parentheses)

	T1		T2		T3		T4	
	I	WC	I	WC	I	WC	I	WC
ULS	52 (100)	50 (100)						
CDI	52 (100)	49 (98)	33 (80)	45 (100)	34 (94)	37 (97)	30 (91)	35 (97)
WellComm			27 (71)	43 (96)	23 (68)	36 (97)	29 (97)	34 (94)
HLEQ			38 (93)	45 (100)	34 (94)	37 (97)	30 (91)	35 (97)
MORS			38 (93)	45 (100)	34 (94)	36 (95)	30 (91)	35 (97)
SDQ			38 (93)	44 (98)	34 (94)	37 (97)	30 (91)	35 (97)

Note. 1= Intervention Group; WC= Waiting Control Group

Table 2 shows the number of completed assessments at each time point for each group. The universal language screener has 100% completion rate in both groups. The CDI completion rate was over 90% in both groups with the exception of T2 for the intervention group. The HLEQ, MORS and SDQ all had completion rates over 90%. In contrast the completion rate for the WellComm was variable and ranged between 68-97%. Completion of this measure was particularly low in the intervention group, with poorest completion at T2.

Reliability

To assess reliability/internal consistency, Cronbach's alphas were computed for all outcome measures from T2 with the exception of the language screener, which was computed from T1 (see Table 3). All of the measures with the exception of the MORS Warmth had reliability figures over 0.75.

Relationship between language and non-language measures.

Initial analysis revealed a significant positive correlation between the language screener and the CDI understanding (r = 0.62, p <.001) and CDI speaking (r = 0.65, p <.001) providing an indication of the

validity of the language screener. To explore the relationship between our language and non-language measures we looked at the correlations at T2 and T4. Table 4 shows that the MORS-Warmth correlated with CDI Understanding and the WellComm at T2, with the HLE correlating with CDI Speaking and the WellComm. At T4, the MORS-Warmth correlated positively with the CDI Speaking, CDI Understanding and the WellComm, while the SDQ correlated negatively with CDI Understanding.

Table 3. Internal reliabilities for the Language screener and non-language measures

	N	Items	Cronbach's alpha
Language screener	95	10	0.77
HLE	82	16	0.76
MORS Warmth	81	7	0.57
MORS Invasive	83	7	0.78
SDQ Total	82	10	0.76

Note. Cronbach's alphas could not be computed for the CDI or WellComm because the data was received as a single score and not at the item level.

Table 4. Correlations between language and non-language measures at T2 and T4 (T4 correlations in parentheses)

	CDI Understanding	CDI Speaking	Wellcomm
HLE	0.2 (0.19)	0.28* (0.06)	0.49*** (0.18)
MORS Warmth	0.23* (0.39**)	0.21 (0.28**)	0.36** (0.30*)
MORS Invasive	-0.05 (-0.16)	-0.13 (-0.12)	-0.25* (-0.01)
SDQ Total	-0.02 (-0.31*)	0.2 (0.21)	-0.26* (-0.20)

Note. *p<.05; **p<.01; p<.001

Group Differences

Descriptive statistics for all measures at all time points are shown in Table 5. For the language measures (Screening score, CDI measures, WellComm), as well as the HLE and MORS Warmth, increases in scores represent improvement, while for the MORS Invasive and the SDQ measures, decreases in scores represent improvement.

Considering the data on the language measures, it is clear that children in the two groups began the trial with similar scores and made progress over time. For the HLE, the mean scores at T2 were similar (although slightly lower in the intervention group), but while the intervention group improved over time, the means for the waiting control remained stable.

Table 5. Descriptive statistics for all assessment measures at all time points for the total sample, the Intervention group, and the Waiting control group

	Full oTTer sample N102						Group A - Intervention						Group B Waiting Control					
	n	mn	sd	mdn	r	CIs	n	mn	sd	mdn	r	CIs	n	mn	sd	mdn	r	CIs
Language Screener (T1)	102	10.85	4.06	11.0	1-19	10.06, 11.64	52	10.48	3.96	10	2-18	9.40, 11.56	50	11.24	4.18	11.5	1-19	10.08, 12.40
CDI Understand																		
T1	101	46.41	27.66	53.0	0-100	41.01, 51.80	52	44.40	27.71	49	0-100	36.87 51.94	49	48.53	27.74	55.0	0-86	40.76, 56.30
T2	78	57.69	23.47	63.0	0-100	52.48, 62.90	33	51.85	26.58	54	0-91	42.78 60.92	45	61.98	20.14	66.0	18-100	56.09, 67.86
T3	71	67.99	19.43	72.0	0-98	63.47, 72.50	34	68.50	19.20	72	0-96	62.05 74.95	37	67.51	19.89	73.0	17-98	61.11, 73.92
T4	65	78.66	15.86	81.0	28-99	74.81, 82.52	30	80.90	15.47	84	28-99	75.36 86.44	35	76.74	16.15	78.0	38-99	71.39, 82.09
CDI Speaking																		
T1	102	21.76	20.92	14.5	0-79	17.70, 25.83	52	19.71	21.23	11.5	0-79	13.9425.48	50	23.90	20.59	21.0	0-72	18.19, 29.61
T2	78	26.78	22.70	18.5	0-82	21.74, 31.82	33	24.67	23.65	16	0-82	16.6032.74	45	28.33	22.12	25.0	1-74	21.87, 34.80
T3	71	38.51	25.70	40.0	0-96	32.53, 44.48	34	41.97	28.98	41	0-96	32.23 51.71	37	35.32	22.19	35.0	1-73	28.18, 42.47

	T4	65	58.89	28.61	64.0	0-99	51.94, 65.85	30	59.67	29.54	63.5	0-99	49.1070.24	35	58.23	28.20	67.0	2-96	48.89, 67.57
WellComm																			
	T2	70	2.14	0.8	2.00	1-4	1.95, 2.33	27	2.19	0.88	2.00	1-4	1.85, 2.52	42	2.12	0.76	2.00	1-4	1.89, 2.34
	T3	59	2.47	0.75	3.00	1-4	2.28, 2.67	23	2.61	0.78	3.00	1-4	2.29, 2.93	36	2.39	0.73	2.00	1-4	2.15, 2.63
	T4	63	3.19	1.03	3.00	1-5	2.94, 3.44	29	3.31	1.00	3.00	1-5	2.95, 3.68	34	3.09	1.06	3.00	1-5	2.73, 3.44
HLE Score																			
	T2	83	29.60	10.55	31.0	1-49	27.33, 31.87	38	28.68	10.74	29.5	8-49	25.2732.10	45	30.38	10.45	32.0	1-49	27.32, 33.43
	T3	71	31.04	10.28	31.0	9-54	28.65, 33.43	34	31.56	10.44	31	13-46	28.05, 35.07	37	30.57	10.26	30.0	9-54	27.26, 33.87
	T4	65	32.17	9.86	34.0	12-52	29.77, 34.57	30	34.13	8.80	36	12-49	30.99 37.28	35	30.49	10.53	31.0	12-52	27.00, 33.97
MORS Warmth																			
	T2	83	28.96	4.69	29.0	15-35	27.96, 29.97	38	28.84	4.47	29	17-35	27.42 30.26	45	29.07	4.91	30.0	15-35	27.63, 30.50
	T3	70	30.26	4.44	31.5	13-35	29.22, 31.30	34	31.00	4.27	32	13-35	29.56 32.44	36	29.56	4.54	30.0	20-35	28.07, 31.04
	T4	65	30.68	4.00	32.0	18-35	29.70, 31.65	30	32.10	2.82	32.5	26-35	31.09 33.11	35	29.46	4.47	31.0	18-35	27.97, 30.94

MORS Invasive																		
	T2	83	11.04	6.19	10.0	1-32	9.71, 12.37	38	10.74	4.97	10.5	2-20	9.16, 12.32	45	11.29	7.10	10.0	1-32	9.21, 13.36
	T3	70	9.97	5.41	10.0	0-24	8.70, 11.24	34	9.88	5.14	9.5	3-20	8.15, 11.61	36	10.06	5.72	10.0	0-24	8.19, 11.92
	T4	65	10.65	5.54	10.0	1-24	9.30, 11.99	30	9.33	4.14	9	1-21	7.85, 10.81	35	11.77	6.36	10.0	2-24	9.67, 13.88
SDQ Score																			
	T2	82	8.13	4.07	8.0	1-18	7.25, 9.01	38	8.05	3.36	8	2-15	6.98, 9.12	44	8.20	4.63	8.0	1-18	6.84, 9.57
	T3	71	7.23	4.36	7.0	0-20	6.21, 8.24	34	6.26	3.53	6	0-15	5.08, 7.45	37	8.11	4.89	8.0	1-20	6.53, 9.68
	T4	65	7.22	4.09	7.0	1-20	6.22, 8.21	30	5.90	2.88	6	1-11	4.87, 6.93	35	8.34	4.65	7.0	1-20	6.80, 9.88

Note. Mn = mean, sd = standard deviation, mdn = median, r = range, CIs = 95% confidence intervals

This pattern of slightly greater progress in the intervention group over the waiting control group was also seen for the MORS Warmth measure. Both the MORS Invasive score and the SDQ total score decreased over time for the intervention group, but remained fairly stable for the waiting control group, suggesting more improvement in the intervention group.

We looked at the difference between groups at T2 and T4 to identify trends in the primary and secondary outcome measures in order to consider which to use in a future definitive trial. Figure 5 shows the mean difference between groups on the language measures CDI Understanding, CDI Speaking and WellComm at T2 and T4 presented as effect sizes with 95% confidence intervals. Each variable is plotted twice, once for the mean difference between groups at T2 and once for the mean difference between groups at T4. An effect size above zero means an advantage for the intervention group. The results indicate a shift in favour of the intervention group over time. For both the WellComm and CDI Understanding, the effect size of the group difference was above 0.2 which is the threshold typically recognised as educationally significant. For CDI Speaking, the effect size at T4 is *d*=0.05 which is not educationally significant. All of the confidence intervals cross zero indicating that none of these differences would be statistically significant. Moreover, the confidence intervals are quite wide, indicating there is a lot of noise in the data. However, given the age of the children, and the lack of stability of language at this age, this is not unexpected.

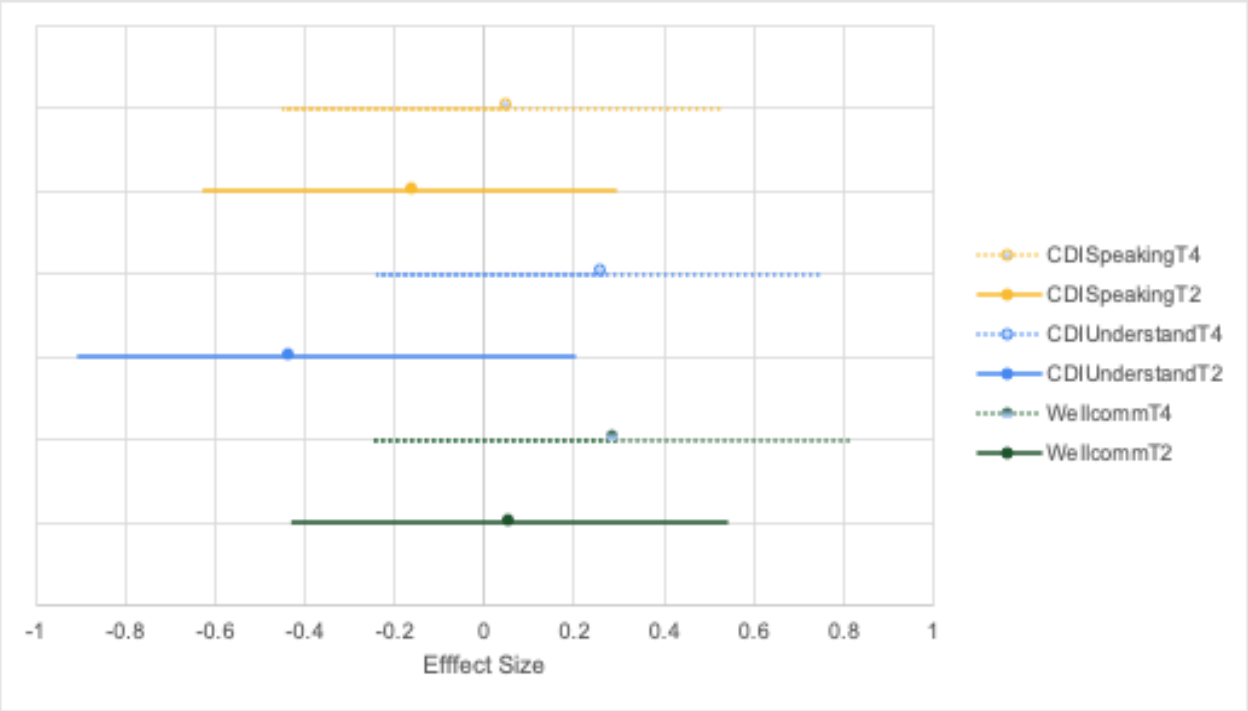

Figure 5. Effect size of mean difference between groups at T2 and T4 and 95% confidence intervals for language measures

Note. Lines of the same colour represent the same variable but at different time points. Solid lines represent T2, dotted lines represent T4. Effect sizes above zero represent an advantage for the intervention group.

Figure 6 shows the mean difference between groups at T2 and T4 for the non-language measures Home Learning Environment (HLE), Strengths and Difficulties Questionnaire (SDQ) and Maternal Object Relations Scale (MORS). The mean difference is represented as an effect size with 95%

confidence intervals. Each variable is plotted twice, once for the mean difference between groups at T2 and once for the mean difference between groups at T4. For the SDQ and MORS-Invasive (MORS-I), an effect size below zero indicates an advantage for the intervention group. For the MORS-W and the HLE, an effect size above zero means an advantage for the intervention group. The results indicate a shift in favour of the intervention group over time. We see a reduction in problematic behaviour as measured by the SDQ and a reduction in MORS-I scores, as well as an increase in HLE and an increase in MORS-W. The confidence intervals are not as broad as the language measures and in the case of the MORS-W, the confidence intervals at T4 do not cross zero indicating that this difference would be statistically significant. All of the T4 effect sizes are above the *d*=0.2 threshold.

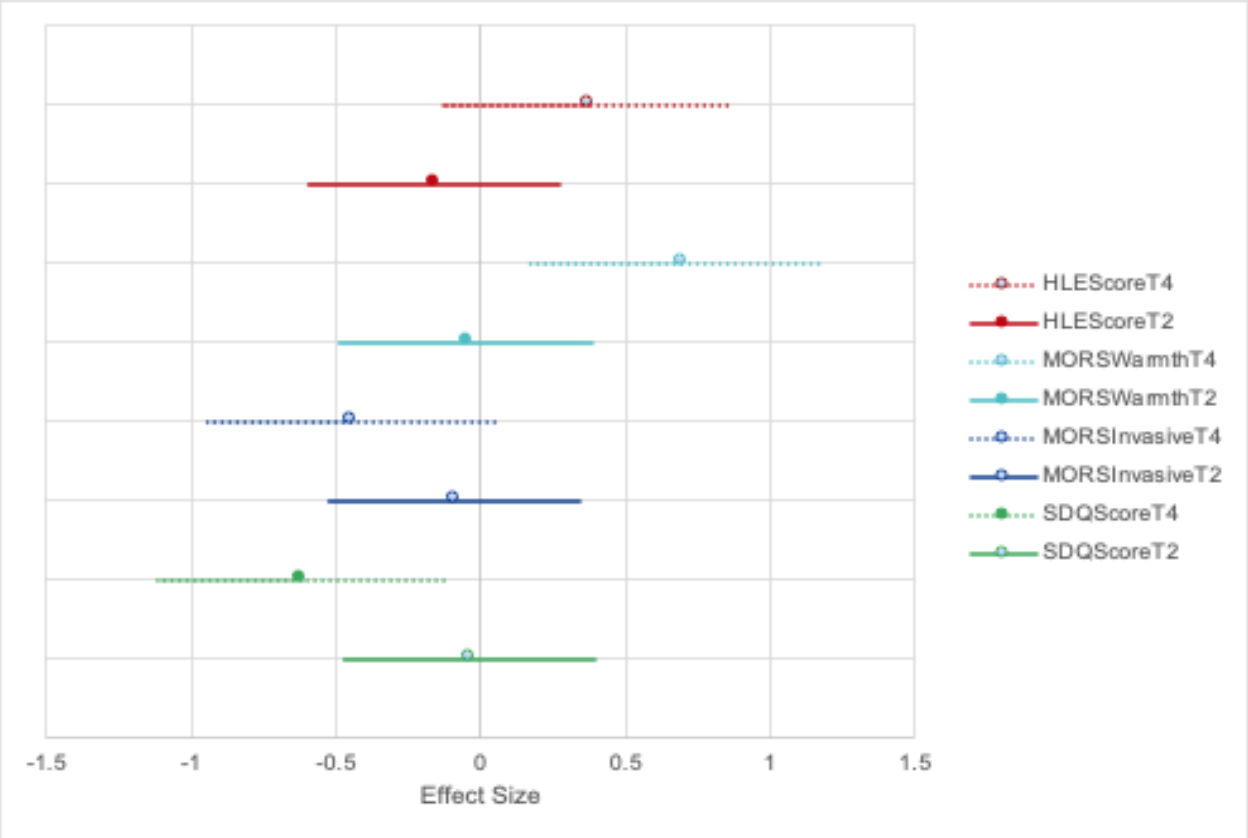

Figure 6. Effect size of mean difference between groups at T2 and T4 and 95% confidence intervals for non-language measures

Note. Lines of the same colour represent the same variable but at different time points. Solid lines represent T2, dotted lines represent T4. For the SDQ and MORS Invasive, effect sizes below zero represent an advantage for the intervention group. For MORS Warmth and HLE, effect sizes above zero represent an advantage for the intervention group.

We also estimated the responsiveness of all measures, calculating the standardised response mean (SRM) for each of the measures by dividing the mean difference of each variable at T4 with the standard deviation of the mean difference. The results are shown in Table 6, the language measures being the most responsive to change over time, particularly the CDI measures, and the MORS-W being the most responsive of the non-language measures, although an SRM of 0.44 is small.

Table 6. Standardised Response Mean of both language and non-language measures

	MeanDiff (T2-T4)	SD of MeanDiff (T2-T4)	SRM
CDI Understanding	33.95	28.20	1.20
CDI Speaking	36.51	30.89	1.18
WellComm	0.95	1.23	0.77
HLE	2.83	11.58	0.24
MORS Warmth	1.90	4.29	0.44
MORS Invasiveness	0.25	5.81	0.04
SDQ total	-0.90	4.56	-0.20

Acceptability of measures

When parents were asked how they felt about the measures, some parents commented that they felt the measures were long and time consuming, and some felt that they took away from the LDW's time working with the child. However, the majority found the assessments to be of an acceptable length or had no real recollection of them, suggesting that they did not find them overburdensome. Some of those parents who found the measures too lengthy noted that the support from LDWs made them more tolerable. Several parents felt that the measures were useful in that they allowed them to understand their child's level of development and track change.

Overall, parents did not seem to have very specific recollections of the measures. The one exception to this was that some parents felt that the Oxford CDI was challenging. A number commented that they felt it was disheartening, particularly when their child knew very few words, and a small number felt that some of the words on the list were too advanced or inappropriate for two-year-old children.

For staff, the additional measures increased their workload, particularly in terms of the amount of paperwork to be completed during the family visits, although some practitioners did say they got used to the additional demands. Some staff were concerned that the level of paperwork changed their relationship with the families. A new system of data entry was also a challenge, both in terms of learning the new system and the time taken to enter the data after each session.

> *"I thought it was a lot of extra work for us to do and although I was, I was really enthusiastic about the study itself and the outcomes of the study for the long-term gain, I felt like I didn't particularly, as a language development worker, want to be having that extra work, having that extra pressure. But once we got started with it, it wasn't that bad, it did take longer."*

Because the additional paperwork had increased the time taken for the initial assessment session the management team had to reorganise workloads to ensure that an LDW did not have two week one sessions on the same day, which resolved LDW concerns regarding the timing.

Some concern was expressed over the measures themselves. For example, one LDW commented that some items in the MORS were awkward to ask parents at first:

> *"one of them was does your child annoy you. I was quite like, ooh, it's a bit of a... You know, a bit of a funny question and I was always quite awkward asking that at first. Now, before it ended, I kind of got used to it because parents actually answered the question"*

Indeed, overall, practitioners noted that their concerns with the assessment measures were often remedied with time or support from the team or management, such as allowing for more time in the sessions with parents, or time for additional data entry.

Identification of most appropriate outcome measures

Taking into consideration both the quantitative performance of the measures, and the qualitative accounts of their acceptability to parents and practitioners, the proposed primary outcome measure for a future trial would be the CDI Understanding. Of the three potential language measures (CDI Understanding, CDI Speaking, and WellComm), the CDI Understanding was considered the strongest candidate due to the consistency with which it was completed at the appropriate timepoints, its strong SRM score, and an advantage for the intervention group on group differences in scores at T4. Although the CDI Speaking performed similarly in terms of data completeness and SRM, it did not show the same evidence of promise when the groups' performances were compared. While it was hoped that the WellComm would perform well due to the benefit of using a holistic assessment of language, it was poorly completed by the LDWs in the intervention sessions. This lack of acceptability and resulting data loss suggested it was not a strong candidate measure for a future trial.

The proposed secondary outcome measure would be the MORS - Warmth. This measure was completed consistently by the LDWs, was considered acceptable, correlated well with the language assessments, and showed the largest group difference favouring the intervention group. All of the other measures (MORS - Invasive, HLEQ, SDQ) showing stronger levels of reliability than the MORS - Warmth, but none were as responsive (SRM), nor did they show as consistent a relationship with the language measures, or as large of a group difference at T4. As such, despite the relatively weak reliability of the MORS - Warmth, it was considered the strongest option for a secondary outcome measure.

> *4. What is the sample size needed for a definitive trial based on data on intervention completion and attrition rates, along with outcome data group differences and variability between study arms?*

Our sample size calculation is based on the CDI Understanding as our primary outcome measure. Educational interventions typically view an effect size of d=0.2 as educationally significant so we have used the observed effect size of d=0.26 in our calculation. We have not controlled for the correlation between CDI Understanding at T2 and T4 because the relationship is weak, which is not unexpected given the lack of stability of language at this age. We have used a one-tailed test to reflect the directional hypothesis. With T4 group means of 80.9 (Intervention) and 76.74 (Control) and a pooled standard deviation of 15.84, and with 90% power to observe our desired effect size we would need a

sample size of 500 (250 per group). To account for observed attrition at 32% the final sample size would be 660 at the recruitment stage (315 per group).

5. Was the intervention delivered with fidelity to the standardised procedures as measured by the frequency and duration of support received by participants?

Having been given the universal language screener (T1), if a family is referred into Talking Together and accepts the programme, they should be seen for their first session (T2) within one month. The intervention comprises six weekly sessions. As such the final intervention session (T3) should be within six weeks of the first session. Due to holidays, illness and cancelled appointments, this isn't always possible, but ideally the maximum time a family should wait between the first and last session is ten weeks. The follow-up session (T4) should be within six months of the first session (T2). For the purposes of the study, waiting control group assessments were aligned with intervention assessments. Table 7 shows the waiting time in weeks for the intervention and control group between each time point. The results in Table 7 show that the average waiting time between assessments for the intervention group were longer than the target waiting times, particularly between T1 and T2, and between T3 and T4, both of which were in red according to the progression criteria. Waiting times for the waiting control group between T1 and T2, and T3 and T4, were longer than target waiting times although these were in amber in terms of progression criteria. Waiting times between T2 and T3 met the target.

Table 7. Waiting time in weeks for participants between each assessment time point

	Intervention			Waiting control		
	T1 to T2	T2 to T3	T3 to T4	T1 to T2	T2 to T3	T3 to T4
Mean (SD)	5.72 (6.54)	8.07 (2.79)	28.32 (3.96)	4.18 (1.87)	6.95 (1.28)	27.29 (4.51)
Range	1.86-43.86	4.86-15.86	22.00-39.29	2.00-11.29	5.00-10.00	22.00-50.00
N Within PC	21	24	19	29	34	25
% Within PC	51	67	58	64	89	69

Note. PC = progression criteria range for waiting time. Red coloured boxes denote that the value of the box is within the red range for the progression criteria requirement, and similarly for amber and green boxes.

6. What are the time and resources required to train practitioners to administer the intervention, and how do these relate to resource requirements for definitive RCT development?

All LDWs undergo extensive training before delivering Talking Together. They receive both ICAN Early Talk training & ELKLAN training - both packages that focus on children's language development. They also complete Makaton training, as well as child safeguarding, awareness of child abuse and neglect, awareness of domestic violence and abuse, and the Bradford protocol for missing children. Because they are working directly with parents in the home they receive training targeted at identification of risk and children's needs and working in partnership with parents. Finally, they receive GDPR training and training in using the data entry system.

In terms of training in the delivery of the programme itself, trainees shadow an LDW delivering one full Talking Together programme, deliver one full Talking Together programme alongside an LDW, and one full Talking Together programme with the LDW observing. In addition, trainees shadow an LDW delivering the two year visit for one month, deliver the two year visit for one month alongside the LDW, and deliver the two year visit for one month with the LDW watching.

Attending all of this training equates to 31.5 full days for the LDW in training, and the same number of days for the member of staff carrying out the training. There are additional charges for the ELKLAN, ICAN and Makaton training courses, and some set-up costs to cover resources and IT equipment, plus costs for travel for the LDWs to visit families at home.

For delivery, each LDW delivers a universal screen session (one hour), six intervention sessions (six hours) and one follow-up session (one hour). An additional three hours is used for administration (arranging appointments, entering data etc.). This equates to approximately 1.5 days of LDW time. LDWs can claim time for travel to and from family homes.

Tables 8a and 8b provide a breakdown of the cost per LDW for training, and the cost per family for delivering the programme. In the oTTer project we had ten LDWs working with 102 families; approximately ten families per LDW. The sample size for an RCT has been estimated at 660 families. We have therefore provided a costing for 66 LDWs to be trained in and deliver the Talking Together programme as part of an RCT. We costed the time for the LDW and trainer at a day rate of £81 per day based on the average salary of a LDW. Travel is calculated on the basis of eight x ten mile round trips (one screening, six intervention and one follow-up) for each family at 40p per mile.

Delivery costs for Talking Together are low at approximately £150 per family (see table 8b). Training packages and resources for set up are also relatively low at £1665 per LDW (see table 8a). The biggest cost is time taken to train practitioners. Cost of trainer time has been calculated on the basis of the existing training model (excluding on-costs). To roll out the programme a new training model may be required and as such these costs may vary. For trainees, where this intervention is integrated into existing services, this time would be incorporated into the existing job description of the person who was taking on the role of LDW and delivering the intervention. This would only be an additional cost if new staff needed to be employed. Moreover, while this upfront cost may be high, once trained, one LDW can deliver the intervention to multiple families making the ongoing cost low. Calculated on a basis of one LDW delivering to ten families, total training cost can be divided by ten to get an estimated unit cost of £676.80. Added to the delivery cost of £153.50 per family, this equates to a total unit cost of £830.30, which compares favourably to other home visiting programmes. For example, Parent Child+ and Parents as First Teachers have both been given a medium to high cost rating by the EIF Guidebook (https://guidebook.eif.org.uk/) which equates to a unit cost between £1k and £2K.

Table 8a. Estimated costs of training for Talking Together for one LDW and corresponding costs for 66 LDWs needed for a sample size of n=660

	#days n=1 LDW	Cost (£)	#days n=66 LDW	Cost (£)
Training Costs				
Trainee LDW	31.5	2551.5	2079	168,399
Trainer	31.5	2551.5	2079	168,399
Training packages	-	665	-	43,890
Resources and IT	-	1000	-	66,000
Total Training costs		***6768***		***446,688***

Table 8b. Estimated costs of delivery of Talking Together to one family and corresponding costs for 660 families

Delivery Costs		***n=1***		***n=660***
Days delivery and admin	1.5	121.50	714	80,190
Travel	-	32	-	21,120
Total Delivery Costs		***153.50***		**101,310**

7. & 8. How acceptable are the intervention and trial procedures for practitioners and families, including randomisation and completion of outcome measures? What are the barriers and facilitators to engagement of families with the intervention and the trial?

Acceptability of intervention and trial procedures, and barriers and facilitators to engagement were addressed during the qualitative interviews with staff and families.

Understanding of the trial and informed consent

Overall, parents seemed unclear on the distinction between the intervention (Talking Together), and the research project to evaluate the intervention (oTTer). When parents were asked about the information they received about the study, they often reported that there had been a lot of paperwork, and that they felt that the project had been explained to them. However, many admitted that they either could not remember the consent process specifically, or they had been distracted at the time. When specifically asked whether they had been told that it was possible to receive the intervention without being part of the trial, the majority of parents reported that they did not think this had been explained to them or they could not remember it being explained to them. Some were even surprised to find that they were part of a research project. Many parents could not recall how they had been introduced to the intervention and research project, with the majority of parents reporting that it was first explained to them by their health visitor (which was not the case). A small number of parents suggested it would have been useful if the distinction between the intervention and research project had been explained more clearly, as this may have impacted on their decision to participate.

Practitioners also noted that some parents did not seem to fully understand the trial. They felt there might have been some parental confusion over oTTer compared to Talking Together, and some parents did not fully understand and therefore just said no to oTTer. They reported four main reasons for the difficulty of understanding for the parents. Firstly, language barriers made it difficult to communicate information clearly to families Secondly, conceptual understanding of research in general. For example, one LDW said;

> *"I think a lot of the parents, you know, they might not have come across anything like this before, they don't fully understand the purpose of, you know, and such a, like research on this scale."*

Thirdly, parental reading ability was also mentioned as a barrier, as some parents were not able to access the information sheets provided. The consenting session included a lot of information, and some practitioners felt parents were overwhelmed by the time they reached the oTTer consent process. For example:

"So, in terms of that, it was quite difficult for some families to still engage, some families had shut down by the time we'd got to the point of offering oTTer."

Finally, some LDWs mentioned busy homes and distraction during the informed consent process as a barrier, something that several parents also commented on as they felt they may have been distracted during the consenting process.

Randomisation and waiting control group

With regards to randomisation, this was also an area of confusion and misunderstanding for parents. Few parents demonstrated an understanding of the purpose of randomisation, and the majority thought that the potential 6-month wait was simply due to a waiting list. However, most parents did report that randomisation and the potential of being asked to wait six months before starting the intervention had been explained to them.

Some parents mentioned that they felt their allocation (to the immediate intervention or waiting control group) could have been explained to them more clearly, and others seemed to be unclear about their group allocation.

When parents were asked directly whether the potential waiting time was a problem for them, most parents said they were happy to wait (13 parents). For example, parents said:

> *"So everything I was kind of expecting to take a long time, because there's waiting lists for everything isn't there, so. It didn't worry me in that sense because I just thought that's just what it's going to be like, it is."*

> *"No, it's alright, you just have to wait your turn."*

However, seven parents stated that it was more beneficial to have the intervention straightaway, while by contrast only two thought it was unlikely to make a difference when children received Talking Together. Five parents had specific reservations about the potential waiting period, with one stating:

> *"Yeah, six months, it's very long time, I didn't want to wait six months. With this age, two/three-year-old, six months, it can change, everything can change."*

By comparison, a few parents specifically mentioned that the young age of the child was why they were happy to wait, either because they felt it was too early to be very worried about their child's language, or because they were happy for the intervention to happen any time before the child began nursery at three years.

Overall, while parents did not demonstrate a good understanding of the reasons for the 6-month waiting control condition, they were aware that this could have been the outcome for their family, and generally they were accepting of this aspect of the project. However, there was a general preference for receiving the intervention immediately.

Practitioners also found that it could be challenging to explain the study and randomisation to parents and ensure they fully understood this component of the project. They commented that they relied heavily on the information sheet, talking through the information using simple words and trying to rephrase where necessary to ensure comprehension. However, sometimes it seemed parents were not clear on what the project entailed, in which case they were not recruited. For example, one practitioner said:

> *"It totally depended on the family to be honest, which I found quite difficult trying to explain it if they didn't fully understand it and sometimes we'd find ourselves trying to explain it you know, two, three sometimes four times and we just sort of had to leave it at that point then because obviously their understanding I felt wasn't enough to be able to sign them up for something that we were asking them to, because I didn't feel that that would have been right if they didn't fully understand."*

Most staff mentioned that the possibility of a six-month wait dissuaded many families from taking part in the trial. Staff felt that this was a significant concern for families, and seemed to be a concern even for those families that decided to consent to the project. One LDW thought that the language packs and visits from the research assistants helped to ease the concern over the wait, although another LDW felt that the packs were not a sufficient incentive for parents to risk the wait. Some practitioners also noted that it was important for them to stress to families that they had no control over the allocations.

Many staff noted they personally found it "difficult", "uncomfortable" or "worrying" not knowing which group a family would be placed in, especially for children with more severe language delay. For example, one practitioner commented:

> *"Sometimes I found it very difficult knowing that some children may have to wait six months and I could see that they needed it immediately or within you know, the very near future."*

This resulted in some practitioners not offering the trial to all eligible families, because they were so concerned about the child waiting to receive the intervention.

> *"but for some, you know, you could clearly see there's a language delay there, it's better just not offering it because you don't want to risk them being put into waiting six months, so I think it was that really. And for some it was inappropriate because you could clearly see that they need it, like, very non-verbal, they wasn't saying anything, so for some times that I didn't offer it because I just thought it wasn't appropriate, yeah."*

By contrast, some LDWs did not have an issue with the waiting list, citing their understanding of the aims and importance of the trial as the main reason.

Another challenge some practitioners encountered was explaining to families in the waiting control group that they needed to continue waiting, especially in cases when families wanted to receive the intervention as soon as possible. This seemed to stem from the issue of parents misunderstanding the 6-month waiting period as merely a waiting list, and as such requesting to be seen more quickly.

Overall, practitioners reported working hard to explain the research project clearly to parents, but were aware that numerous parents simply did not understand what the project entailed. The data from practitioners also makes it clear that the waiting control group was a real issue for many parents, more so than the family data because those parents would have had to be okay with the wait in order to agree to the project. They also found it difficult themselves to deal with the waiting period, particularly for those families they considered most in need.

Barriers to the intervention

Parents did not report many barriers to receiving the intervention. A small number of parents reported that they were not necessarily as concerned about their child's language development as the LDWs, and felt that it may be worth waiting longer before suggesting there was a delay. This was also reflected in some parents' comments that the LDWs' assessments of their child as being behind in their language development caused them to worry (in some cases they felt this worry was unnecessary).

Some parents found their child's behaviour during sessions to be challenging, with mentions of their distractible or disruptive behaviour. One parent commented that her child's behaviour was only difficult when there were visitors in the home, and this made the sessions stressful for her.

The most consistent comment about the intervention was that the sessions were considered too short. Several parents mentioned that they would have liked the sessions to be extended to allow more time to cover the content, or to deal with distractions.

Practitioners also reported that some parents did not want the intervention because they did not feel their child had a problem, even if this was contradicted by the assessment. The staff reported that some parents thought that their child would catch up with time, and this also applied in some cases where the children were learning multiple languages.

LDWs mentioned that parents were often busy with multiple commitments, and found it difficult to fit in appointments around other activities. This was also related to challenges with getting families to engage with the full programme. For example, one practitioner stated:

> *"But yeah, it was trying to get in and making them understand that it needs to be consistent, that's why we need to come in for the full six weeks and not, oh yeah, sorry I forgot that you were coming today, I'm in town, you know and just sort of like ringing up and oh I'm not going to be in today, I'm going out and that was sort of like things that I come up against and I'd say to them, it's really important though that I do come every single week because it's got to be consistent for this to be successful for your child."*

According to the practitioners, some families were very difficult to engage, and never really fully committed to the intervention. This resulted in them never starting the intervention, or skipping or rearranging sessions to the point that it is not possible to deliver all of the course content. One staff member suggested it could be social desirability that caused parents to accept Talking Together at the screening visit but then drop out before the first session:

> *"But we get drop-out in Talking Together as well once parents start to, you know, do the programme [...] do parents actually know what they're signing up for at the point that they're signing up for that, and do parents say yes to Talking Together because they've just been informed that their child has a language delay and they think they need to do something about it? Or do they sign up to Talking Together because they actually want to do something about it?"*

Facilitators to the intervention

An important facilitator that several parents mentioned was the home-visiting nature of the intervention. Parents appreciated that the LDWs came out to families' homes, stating that this was easier and more convenient for them. Several parents also noted that they liked that the intervention is one-to-one, perhaps in contrast to group-based sessions. These core elements of the intervention design were clearly appreciated by parents.

Another central facilitator to Talking Together was the relationship between LDWs and both children and parents. LDWs' training highlights the importance of developing good, trusting relationships with families, and the team's success in this area was clear from the comments parents made about the staff. For example:

> *"Yeah, he used to love it, like when she was coming on the front he was like [...] he was constantly in the window looking for her, so he knew like it was, you know, like a fun time when she came really."*

Parents spoke of the bonds their child had made with the LDWs, how their child enjoyed the sessions, and how they looked forward to LDWs arriving. Many parents noted that they were pleased with how the LDWs engaged with their child, and felt they were professional and polite. They also developed positive relationships with the parents themselves, for example:

"...you look forward to the session, it was just good, she made me feel so comfortable, it's like talking to my friend, you know when it wasn't very like strict in the sense that I felt like she gave me the chance where I can talk about how I feel, any concerns I have she made me comfortable to be able to speak..."

Parents also felt like the appointments were convenient for them, and that LDWs were helpful and accommodating of their schedules.

Finally, while not all parents thought that the intervention was going to make a difference to their child, several spoke about hoping that Talking Together would improve their child's language skills. They noted that concern for their child's language development and the potential that the programme could support their child in this skill prompted their decision to engage with Talking Together.

Practitioners noted several of the same facilitators to the intervention. Akin to parents, they felt that the home-visiting nature of the project was a facilitator to uptake and supported parents to continue through the intervention. Similarly, staff commented on the importance of building trusting relationships with families and children. Some staff noted that this focus on building trusting relationships is a fundamental component of LDW training, along with cultivating LDWs positive approach to families.

Staff also mentioned that they worked hard to accommodate parents and to support them how they could, and the parent feedback demonstrated that their efforts were felt by participants.

Finally, staff mentioned the importance of access to interpreters to ensure that the programme was accessible to the whole community, regardless of home language.

Barriers to the trial

The most notable barriers to the trial for parents have all previously been discussed. Firstly, parents' concern about being allocated to the waiting control group was a large barrier to the trial, and resulted in some parents choosing not to consent to be involved. Secondly, parents struggled to understand the difference between the intervention and the research. This is an important concern, as it is vital that parents fully understand their right to receive the intervention without being involved in the research. Finally, although parents were generally accepting of the assessment measures, the time taken to complete these assessments, and parents concerns about the appropriateness of the CDI should be considered a barrier to the trial.

Additionally, one minor challenge for the trial was the need for the intervention to be delivered to a named parent or carer. In normal practice, the intervention can be delivered more flexibly and sometimes different parents or carers within the household are present at the sessions. This supports parents to work around their other commitments while still receiving the intervention, and a parent noted that this would have been preferable during their experience of the trial.

Practitioners also felt there were a range of different barriers to the trial. As noted previously, the challenge of explaining the intervention adequately to parents was a significant barrier for the trial, and would need to be addressed in any future work. Also, practitioners' concerns about randomisation, and their decisions to at times not offer the intervention to eligible families because they wanted to avoid the family being placed in the waiting control group was a major issue for the trial.

However, an interview with a member of the BHT management team suggested that while some eligible families were not offered the trial, most eligible families were invited to take part.

"I wouldn't say that it was a major issue, it happened at points throughout, but not like the five families a week, just one sort of, just once or twice every so often."

There was also some confusion around the eligibility criteria, and the recording of eligibility using the checklist.

"a staff member had not offered it because of a confusion with one of the eligibility criteria's about, I think it was the language, because the parent has to speak it to their child as well as speaking in, parent has to speak English to their child as well and I think that confused one or two staff to begin with"

In addition to these challenges, many practitioners discussed the additional paperwork and new data entry procedures as a considerable burden to them. The impact was felt both in terms of how much time the new procedures took in their sessions with families, as well as the time it took them to enter the data onto the database. Although some practitioners were concerned that the time spent on the paperwork jeopardised their relationships with families, many of the families noted that they had very positive relationships with the LDWs.

Finally, one element of the design of the trial was that research assistants accompanied LDWs on visits while they were being trained and again for all the 6-month reviews. Some LDWs felt that this was sometimes too much for the family. For example, where a child had additional needs, or where there were space and time constraints.

"it would be too overwhelming for him to have two people going into the home to play with him, because it was too much for him.[…]. so, the session was a bit difficult then to get him to engage because there was two of us in the room with mum and it was too overwhelming for him."

Additionally, some LDWs felt that the need to organise the final assessment sessions with the RAs made it more challenging to schedule appointments with families.

"then we're trying to do the reviews, and we need to take an RA with us, and it's trying to fit in those times, because it is difficult, because I know that the RAs have got a lot of their own stuff to do as well as we're like, we need someone to come out, or they're trying to find one of us to go out with them on their last review visits, and the timetables do clash quite a bit, so it has been a bit difficult."

Facilitators to the trial

Parents mentioned few facilitators to the trial over and above the facilitators they mentioned for the intervention (i.e. home visiting, LDW's relationships with families). However, one trial specific facilitator some parents noted was the benefit of having multiple staff members attending the sessions. Some parents felt this made the sessions easier, as one practitioner could engage the child while the other worked with the parent. This was particularly likely to happen for trial families during the assessment sessions (most notably T4, the follow-up session).

Practitioners noted a number of different facilitators to the trial. Some LDWs said that the information leaflet was beneficial in ensuring they covered all the necessary information with families. Like parents, some staff also mentioned the benefit of conducting the visits in pairs, although their reasoning was also about the benefit of teamwork for learning the new trial procedures. Working with the RAs was mentioned positively by some LDWs, and staff noted that they felt they felt they were well trained and supported in their new roles:

> *"we all had lots of training and we all had lots of meetings, we all had lots of discussions so that everybody was clear on what we were doing, and if anybody wasn't clear you know, there's always been somebody there just to clarify. So, I think in that sense, it has been managed quite well, that we've all been fully informed about the process of things and how to do things, we've all been fully supported as well"*

Finally, some practitioners felt that the RA visits and language packs were helpful in encouraging parents to engage with the trial despite their concerns about the waiting list (although not all LDWs felt this was the case).

Impact of the intervention and trial

The majority of parents stated that the programme was beneficial for their child's development. In addition, parent's used adjectives such as "fun" and "really enjoyable'' when describing the programme. Parents also noted the benefits that the programme had for themselves:

> *"I felt like I was learning more, maybe more than he was, sort of thing, it wasn't, it was as much learning for me as it was for him. You know, certainly with non-verbal communication"*

However, two parents stated they did not feel the programme helped, that their child's development was normal, and the parent did not learn anything new. Despite this, all parents would recommend Talking Together to family or friends, and a few had already told family and friends about the programme.

Practitioners also felt the programme had a big impact on children, and many were passionate about the impact the programme had on families in the community. In addition to this, many staff also spoke of being part of the research trial as a positive experience in which they learnt a lot:

"I think it boosted my confidence quite a lot. I actually enjoyed taking part in the project, I thought it was... It gave me an insight into what actually happens when you do research and to be honest, I'd love to know what happens at the end of the research."

BiBBS data

1161 Talking Together families were also involved in the BiBBS cohort study, these families were all referred for language screening between 20/06/2017 and 29/05/2020. 862 families completed the language screening and were either a) not offered Talking Together (385 families), b) offered and accepted Talking Together (349 families), or c) offered and declined Talking Together (84 families). For 44 families we did not have data on their screening outcome.

We examined the BiBBS data for predictors of 1) referral into Talking Together, 2) uptake of Talking Together offers, and 3) outcome of Talking Together. We identified 3 predictors to be used in regression analyses, 1) English as a first language, 2) ethnicity and 3) education. Maternal mental health and socioeconomic status measures do exist in the BiBBS dataset and were considered for our analyses, however due to data completeness we decided not to include them. Descriptive data for each of the predictors for all of the families that completed screening can be found in Table 9.

The predictor English as a first language was a categorical variable with two levels, Yes or No. The categorical variable ethnicity comprised four levels, white British, white other, south Asian, and other. Maternal education was treated as a continuous variable on a 5-point scale, ranging from 1 = *no qualifications*, 2 = *5 or less GCSE (grades A-C) or equivalent*, 3 = *5 or more GCSE (grades A-C) or equivalent*, 4 = *A levels or equivalent*, 5 = *Degree or equivalent*.

Table 9. Descriptives for the Predictors by Talking Together Offer

		BiBBS Sample	Talking Together Offer		
			Not Offered	Offered & Accepted	Offered & Declined
	N	1161	385	349	84
English First Language (%)	Yes	56.10	56.06	58.99	44.44
	No	43.90	43.94	41.01	55.56
	Missing	8.18	7.79	9.17	3.57
Ethnicity (%)	White British	10.62	9.46	10.75	9.64
	White Other	7.29	5.95	4.18	8.43
	South Asian	72.01	74.05	74.33	75.90
	Other	10.08	10.54	10.75	6.02

	Missing	4.31	0.39	4.01	1.19
Maternal Education (%)	No qualifications	9.30	8.02	8.74	14.67
	5 or less GCSE (grades A-C) or equivalent	32.60	30.09	30.42	40.00
	5 or more GCSE (grades A-C) or equivalent	14.19	14.04	13.27	10.67
	A levels or equivalent	12.75	11.46	16.18	6.67
	Degree or equivalent	31.16	36.39	31.39	28.00
	Missing	0.10	9.35	11.46	10.71

Statistical Analysis

Initially, the three predictors 1) English as a first language, 2) ethnicity and 3) education were entered into a logistical regression model to predict the binomial outcome of referral (i.e. families that were referred into Talking Together and families that were not referred). We compared this model to a second model, which comprised only two predictors, English as a first language and education. The two models had similar Akaike Information Criterion's (AIC's), 986 and 987 for models one and two respectively. It was not possible to perform a likelihood ratio test because the models were not all fitted to the same size of dataset (i.e. there was more missing data in the first model). Therefore, the simplest model, model two, was chosen for analysis.

Two logistical regression models were also built to assess uptake of the Talking Together offer (i.e. if the offer was accepted or declined). The first model included the 3 predictors English as a first language, ethnicity, and education, whereas the second model only included English as a first language and education as predictors. The likelihood ratio test (LRT) was used to compare the two models. The LRT was non-significant meaning there was no significant difference in model fit between the two models. In addition, the AIC for the first and second models were 377 and 376 respectively, suggesting that both models were of similar quality. Therefore, the second more parsimonious model was chosen.

Finally we created two logistical regression models to test predictors of Talking Together outcome (i.e. whether the course was completed or not completed once an offer had been accepted). Like before, the first model included the three predictors English as a first language, ethnicity, and education, and the second model only included English as a first language and education as predictors. The LRT was non-significant meaning there was no significant difference in model fit between the two models. In addition, the AIC for the first and second models was 314 and 309 respectively, suggesting that both models were of similar quality. Therefore, the second more parsimonious model was chosen.

Results

According to our analysis, referral into Talking Together was not predicted by either English as a first language or education (Table 10). The 95% confidence intervals for both predictors cross one, this implies there is no difference between levels of the predictors and the dependent variable referral into Talking Together. For example, the odds ratio for English as a first language was 1.02, which suggests that a family has equal odds of being offered Talking Together whether they speak English as a first language, or they do not speak English as their first language. Likewise, the level of education does not influence the odds of being offered the Talking Together programme. It is worth noting that McFadden's R^2, which was used to evaluate the model fit, was very low. This suggests that the model has limited predictive power and as such we can't conclusively accept the null hypothesis that there is no relation between these predictors and referral into Talking Together.

Table 10. Logistic Regression for Referral into Talking Together

	B	*SE B*	*OR*	95% CI
Intercept	0.15	0.20	1.16	0.78, 1.71
English first language Yes	0.02	0.16	1.02	0.75, 1.39
Education	-0.02	0.05	0.98	0.88, 1.09
McFadden's R^2	0.00			

Note. 109 observations deleted due to missing data. Reference categories; referral into Talking Together = "not offered Talking Together", English as a first language = "No". B = beta coefficient. SE B = standard error of the beta coefficient. OR = OR = odds ratio. CI = confidence interval. * $p < .05$

We did find a relation between the uptake of Talking Together and English as a first language (Table 11). The negative beta coefficient of -0.56 suggests that declining Talking Together is less likely to occur for families with English as their first language compared to families that do not speak English as their first language. In other words, a family that does not have English as their first language is more likely to decline Talking Together, and a family with English as their first language is more likely to accept Talking Together. The odds ratio of 0.57 suggests that English speakers have roughly half the odds of declining Talking Together compared to non-English speakers.

We found no relation between education and uptake of Talking Together. Because the 95% confidence interval for education crosses 1 this implies that with varying levels of education there are equal odds of accepting or declining Talking Together. McFadden's R^2 was used to evaluate the model fit, this pseudo R^2 metric is similar to R^2 commonly used in linear regression models. Because McFadden's R^2 was low, this suggests that the model has limited predictive power and therefore the results should be interpreted with caution.

Table 11. Logistic Regression for Uptake of Talking Together

	B	SE B	OR	95% CI
Intercept	-0.79*	0.32	0.46	0.24, 0.84
English first language Yes	-0.56*	0.26	0.57	0.34, 0.96
Education	-0.08	0.09	0.92	0.77, 1.10
McFadden's R^2	0.02			

Note. 61 observations deleted due to missingness. Reference categories; uptake of Talking Together = accepted Talking Together, English as a first language = No. B = beta coefficient. SE B = standard error of the beta coefficient. OR = odds ratio. CI = confidence interval. * $p < .05$

We did not find a significant relation between either of the predictors and the outcome of Talking Together (Table 12). For example, the odds ratio for education was 0.87 which suggests that the odds of not completing the course is slightly less for families with higher maternal education compared to those with less education. However, because the 95% confidence interval encompasses one, we cannot be certain that the true value is not one, in which case there would be no difference between education level and Talking Together outcome. In addition, McFadden's R^2 was extremely low which suggests that this model has minimal predictive power.

Table 12. Logistic Regression for Outcome of Talking Together

	B	*SE B*	*OR*	95% CI
Intercept	0.42	0.36	1.52	0.75, 3.09
English first language Yes	0.07	0.28	1.07	0.61, 1.87
Education	-0.14	0.10	0.87	0.71, 1.05
McFadden's R^2	0.00			

Note. 129 observations deleted due to missingness. Reference categories; outcome of Talking Together = completed course, English as a first language = No. B = beta coefficient. SE B = standard error of the beta coefficient. OR = odds ratio. CI = confidence interval. * $p < .05$

In summary, we found that English as a first language predicted uptake of Talking Together. However, no other associations were found between our predictors and our three dependent variables, referral, uptake and outcome of Talking Together.

Evidence of Promise

The final aim of this study was to consider whether the Talking Together intervention shows evidence of promise, suggesting that it could make a meaningful difference in improving children's language and communication skills over time. This assessment was framed by the intervention's theory of change (see Figure 1). Specifically, the measures included in this study assessed the home learning environment and parenting sensitivity and warmth, as well as children's language and behaviour. The theory of change posits that improvements at the parent level will lead to impacts on children's language and behaviour. Based on the consideration of group differences in favour of the intervention group on the outcome measures, the strongest evidence suggested a route of change through improvements in parental warmth (MORS Warmth) that could be the driving force behind the increases in children's receptive language skills (CDI Understanding). The evidence for intervention specific improvements in the home learning environment and children's behaviour were not as strong, although improvements were also seen on both of these measures. Overall, the evidence from the outcome measures supports the theory of change and provides preliminary evidence of promise for the intervention.

It was also relevant to consider the qualitative results in making a judgement of evidence of promise. Parental interviews revealed that parents were overwhelmingly positive about the intervention, and the majority of parents considered that the intervention had made a positive impact on either themselves, their child, or both. While these anecdotal accounts do not carry the same weight of evidence as the quantitative group comparisons on the outcome measures, it is important to state that this is a well received, widely accepted intervention that many parents feel is helpful. This further strengths the argument that this intervention shows evidence of promise, and should be considered for further, more conclusive evaluation.

Key Findings

The aims of this study were to establish the feasibility of a full-scale RCT of the Talking Together programme and to understand factors in referral, uptake and attrition figures associated with Talking Together which may inform the future provision of the service.

Progression Criteria

The results of our feasibility study suggest that the Talking Together programme has the potential to be scaled up to a full trial, but with some caveats. Results from our progression criteria focusing on recruitment, adherence, and attrition highlighted areas of strength and weakness in the study, as well as potential necessary changes to make a future trial successful.

Recruitment

In terms of recruitment, 74% of families who accepted their referral into Talking Together were eligible for the trial, while 62% of families consented to take part in the trial, which put the project in green for both of these elements of the progression criteria.

Although these results were very positive, there were a number of issues that would be relevant to consider in a future trial. Firstly, the qualitative work revealed that LDWs did not offer the trial to all eligible families. For some, this was a result of confusion over the eligibility criteria. For others, it was a conscious decision not to offer the trial, as they did not think it was appropriate or did not want families to be allocated to the waiting control group. There were also a number of eligibility violations that were discovered at various points throughout the trial, including very late in process. This suggests a number of important changes would be necessary in future, including amending the eligibility assessment process and paperwork to make it clearer, ensuring all staff recruiting to the programme are confident assessing eligibility, and implementing a review procedure to ensure eligibility is closely monitored and all eligible participants are invited.

The eligibility criteria also restricted participants to those who spoke English, Urdu, or Punjabi as the primary languages in the home. However, it was also communicated to LDWs that the reason for this criterion was the need for the intervention to be delivered by a LDW directly (not through an interpreter), and as such families who spoke a language also spoken by a member of the LDW team could potentially participate in the study. This resulted in 11 participants who spoke languages other than English, Urdu, or Punjabi as their primary language being included in the study. These participants were not technically considered as eligibility violations, as the spirit of the language eligibility requirement was about ensuring that LDWs were able to deliver the intervention and assess children appropriately, but this does highlight the importance of being much more explicit about eligibility. In future, it would be important to consider the languages of the communities being served by the intervention to identify consistent eligibility criteria across multiple sites. This is no small challenge, given the linguistic diversity of many communities, as well as the need to ensure that practitioners in the administering teams are multilingual in the appropriate languages. However, while practically it would be most straightforward to limit inclusion to only those families who speak English at home, this would exclude important parts of the communities this intervention could serve.

Another important consideration was the representativeness of those participants recruited to the trial. The demographic details of the two arms of the study were relatively similar, suggesting that randomisation in this sample was successful in creating equivalent groups on the variables measured. The exception to this was gender, for which the waiting control group ended up with a notably larger sample of female participants. Given that there is evidence that boys are overrepresented in samples of children with early language delay, even at two years[27], it may be beneficial in future trials to include gender as a stratification variable. Language background was well balanced across the groups, and this is likely to continue to be an important stratification variable in future work. Indeed, the results from the BiBBS analyses in this report also confirm the relevance of including English as a First Language as a stratification variable, given the potential differential uptake of, as well as engagement with, this project based on this participant characteristic. However, while these two variables (gender, language background) should certainly be taken forward as stratification variables in a future study, it is likely

that this list is not exhaustive. It would be important to consider the specific demographics of the communities receiving the intervention to ensure that any other key variables that have shown evidence of impacting early language outcomes are considered when selecting stratification variables.

Adherence

In terms of adherence, for the intervention group the percentage of families seen within the required time frames ranged from 51% (red) to 67% (amber). For the waiting control group the percentage of families seen within the required time frames ranged from 64% (amber) to 89% (green). The difference between these two groups can likely be attributed to the fact that the intervention group assessments were carried out by LDWs who were dealing with live caseloads, while the waiting control group assessments were carried out by RAs who did not have additional caseloads. This result suggests that it was challenging for the team to adhere to an assessment schedule that was necessary to ensure comparability between the two arms of the study. It was particularly important that the T2 and T4 assessment points lined up, and these time points had the lowest level of adherence to the protocol. The T2 time point was very soon after families were recruited, and given the administrative processes involved with recruiting a participant (entering participant onto the system, randomisation, case allocation to a LDW), as well as the need for the participant to have time to be seen on short notice, it is perhaps not unsurprising that this was challenging. This suggests it would be beneficial to review how the administrative processes associated with recruitment could be streamlined or made more efficient, in order to ensure that participants are contacted as soon as possible to allow more flexibility to see them on their schedule, while meeting the protocol time requirement. For the T4 time point, the qualitative data suggested that this assessment session put particular pressure on the team due to the need for both an LDW and an RA to be present. It may be necessary in future to consider more carefully whether there is sufficient capacity in both teams to ensure that the total caseload can be visited on time.

Attrition

In terms of attrition, our final sample at T4 was 68% of our original sample; an attrition rate of 32%, which puts the trial in red for this progression criteria. The biggest attrition rate occurred between T1 and T2, when 11 participants withdrew from the intervention group and five from the waiting control. This greater loss to the intervention group suggests that attrition was not simply a case of participants withdrawing once they learned of their allocation to the waiting control group in order to avoid the six-month wait. Instead, the qualitative evidence suggested that some attrition may have been due to families not having the time to engage fully with the service, or feeling obliged to sign up when they were offered the programme, but then changing their minds. The qualitative work also highlighted that even some of the families who participated in the study were not clear on what they had consented to be involved in when they agreed to take part. It is possible that some families agreed to take part, but then withdrew once they were contacted to arrange their first assessment and learned more about the study. This is further evidence of the need to reform the consent procedure (as discussed below).

While this attrition rate is certainly suboptimal, it is worth considering the attrition rate of the other interventions running within the community as part of BSB's programme. Overall, for 11 of BSB's other projects with recent and available data, attrition between enrolment and completion ranged from 14-60% with an average of 36%. The overall attrition rate for participants of Talking Together in its most recent contract phase (including all participants of the intervention, regardless of their participation

in this trial) was 43%. As such, participants of the trial had a higher completion rate than the intervention in general, and the attrition rate of the trial was in line with the average rate for interventions in this community. As such, while this is certainly an area that would need to be addressed in future, it is worth keeping in mind that most interventions in this community struggle with attrition. Perhaps the expectations for a future trial should take this high level of attrition into account by design. Indeed, the attrition rate from this study was used in the sample size calculation, thereby setting out a target sample size that would hopefully offset attrition.

Nonetheless, it would still be useful to identify and implement retention strategies, in addition to improving the consent procedure, in a future full trial. For example, both groups would likely benefit from receiving more explicit information about their group allocation and what to expect from the trial at the point of randomisation. Also, attrition was slightly higher in the intervention than control group, which could suggest either that the intervention requirements (e.g. six weekly sessions) were too much for some participants to commit to, or that the practitioners could do more to maintain contact with participants (particularly those at risk of disengagement). In terms of the burden to participants, the results of this study have identified key assessments that could be maintained while others could be discontinued, which would decrease the assessment burden. This in turn could allow more time for participants to develop relationships with the LDW working with them, and would also decrease the burden on practitioners in terms of data entry. The additional capacity could be allocated to ensuring that participant engagement is proactively tracked. This was already a component of this study, but it could be prioritised in future work. Additionally, the waiting control group would probably benefit from additional communication with the RAs, as the qualitative data revealed that some participants felt disconnected and would have appreciated more communication and clarity on what was happening. Finally, it may be possible to allow participants more flexibility in terms of the timeframe for intervention delivery, although this would only improve intervention (not trial) retention unless a different type of control group was used (i.e. a no treatment control).

Primary and Secondary Outcome Measures

We also identified primary and secondary outcome measures for a future trial. Combining the quantitative and qualitative results we would recommend our primary outcome measure to be the CDI Understanding. The measure was consistently delivered at each time point, had a strong SRM, and mean group differences showed evidence of promise in terms of an advantage for the intervention group at T4. While the CDI Speaking measures also had a large SRM and was relatively complete, it did not show much evidence of improvement in terms of group differences. However, we would still include this measure as it may provide useful insight into discrepancies between receptive and expressive language. Finally, the WellComm measure showed evidence of promise in terms of group differences and relatively large SRM, but was not consistently completed by the LDWs. The qualitative interviews indicate that time is a big concern for the LDWs and as the WellComm can take a while to complete it is not appropriate for a larger trial. For our secondary outcome measure we would recommend the MORS-Warmth. Although the reliability is relatively weak, it is the most responsive of the non-language measures, and correlates well with the language measures, which has also been found in previous research[28, 29]. Some authors suggest that maternal warmth creates an environment that enhances turn taking during parent-child interactions and thus may aid language development[30].

In addition, the MORS-Warmth was consistently administered in the trial and showed the most evidence of promise in terms of group differences at the end of the trial. Together these measures capture the language and non-language effects of the intervention. The Home Learning Environment Questionnaire and the SDQ would be removed from the test battery for a full-scale trial.

Additional lessons learned

The qualitative work revealed a number of important barriers and facilitators to the intervention and the trial that were useful lessons for how to optimise a future trial. Practitioners' were concerned that parents struggled to understand the distinction between the intervention and the research, and this was confirmed by the interview data from parents. Although parents reported that the trial had been explained to them, some were unaware that they could receive the intervention without being part of the research, suggesting that their consent was not fully informed. This is an important finding and suggests it would be necessary to change the consent procedure in future to ensure participants have a full understanding of the study. A real challenge for this study was the amount of paperwork parents were asked to complete during recruitment. Recruitment to this trial came at the end of numerous other consent processes, including consenting to data sharing with both BSB and BHT Early Education and Training. LDWs noted in their interviews that parents were often fatigued by the time the information on this trial came around, and so a suggestion for future would be to either decrease or streamline the amount of information parents receive at recruitment. Additionally, it may be useful to allow a period of time for parents to consider the trial before agreeing, or to confirm that parents are still happy to be involved after a short waiting period and before randomisation. These steps could help to ensure parents are fully informed about the study, and may also decrease the attrition we encountered between T1 and T2.

Although the randomisation procedures in this study worked well, the six-month waiting control group was a concern for both parents and practitioners. Many parents noted the potential of being allocated to the waiting control group as their reason for not consenting to the trial, and even some of those parents that consented to the trial felt being part of the immediate intervention group was preferable. The six-month wait also impacted on how and when some practitioners offered the intervention, potentially creating a selected sample based on the concerns of the LDWs for some eligible children. Other barriers included the amount of assessments carried out in the sessions, and the data entry time for the practitioners. This resulted in the suggestion from both practitioners and parents that it would have been preferable to have slightly longer sessions, although the implications of this change for service delivery are unclear.

However, there were also a number of useful facilitators for the intervention and trial. Parents appreciated the home-visiting and one-to-one nature of the intervention, and the convenience of LDWs accommodating their schedules when organising visits in their home. Despite concerns that the assessments would take away from the strong relationships LDWs develop with families, parents reported that both they and their children enjoyed the sessions and felt supported by the LDWs. This is important, as the service sees these trusting relationships as a central component of the intervention and its success. Finally, receiving sufficient training and support from managers and the team were key facilitators to the trial delivery for LDWs, and it would be important in any future trial to consider how

to ensure that practitioners had the opportunities for teamwork and support that they had in this project.

Predictors of Referral, Uptake and Outcome

We analysed BiBBS data to test whether English as a first language or maternal education predicted referral, uptake or outcome to Talking Together. We found no associations between either predictor and referral into Talking Together. This suggests that the programme is offered to all families that are identified as having a need at screening, regardless of their educational level or first language being English.

We did find that English as a first language predicted uptake of Talking Together. Families that spoke English as a first language had higher odds of accepting Talking Together compared to those families who did not speak English as a first language. This finding suggests that an English speaking family would be more likely to accept the intervention than a non-English speaking family. There was no evidence to suggest that mothers' educational level influenced the chance of a family accepting the intervention. Regarding outcome of the Talking Together programme (i.e. whether the course was completed or not completed), neither English as a first language nor maternal education predicted course completion.

The impact of greater uptake of Talking Together by English speaking families is that non-English speaking families are potentially underserved by this intervention. Children from different language backgrounds are equally likely to need support for their early language development, so it is important to ensure that this intervention is as widely accessible as possible. For example, it would be useful to continue to recruit LDWs who speak as many of the community languages as possible, as this is the best option for ensuring that the intervention can be delivered with fidelity to families from non-English backgrounds. Alternatively, interpreters can be used, although it is unclear exactly what the impact of receiving the intervention through an interpreter is for fidelity and participant outcomes. For the trial, it would be important to balance the need for methodological rigour in terms of consistency of intervention delivery and validity of assessment measures, with the need to ensure that the sample in the trial is representative of the population receiving the intervention. This may be challenging, and would likely require multilingual research assistants or assessors.

Evidence of promise

Evidence of promise was considered based on group difference on the outcome measures, as well as parents' experiences of the intervention. Group differences on the outcome measures suggested a potential pathway of change through improvements in parental warmth with impacts on children's receptive language, although there was also some evidence of improvements in the home learning environment and children's behaviour. Many parents reported that they felt the intervention had a positive impact on themselves, their child, or both. Additionally, the intervention was considered to be well received and accepted by this population. As such, the final conclusion is that this intervention demonstrates evidence of promise, and is worthy of additional, more conclusive, evaluation.

Limitations

Although care was taken to ensure methodological rigour in this study, there were a number of important limitations to note. Firstly, the nature of the Big Lottery funding meant that the only control group that could be considered was a waiting control. While this has the strength of allowing for both arms of the study to receive the intervention, its weaknesses were in the need to use RAs to collect the data in the control arm, and the limited follow up time. On this first point, it is was necessary to have RAs collect the data in the control arm, as this was beyond the capacity of the intervention team. While these RAs were carefully trained and worked alongside the LDWs throughout the trial, it is possible that they may have interacted with and assessed participants in different ways to the LDWs, introducing bias to the results. It would have been preferable to have independent assessors (ideally blinded) to assess both groups at either all or key time points (T2 and T4), but this also has practical challenges, especially when attempting to assess language in children who are unfamiliar with the assessor.

On the second point, a six month follow up was all that was considered ethically and practically possible in this study, and was also in line with current intervention practice. However, as children's language skills develop quickly and often unpredictably at this age, a longer follow-up time would be preferable. An alternative or no-treatment control would allow for children's outcomes at later time points to be compared to consider whether there were greater or lesser benefits to the intervention over time.

Another limitation was the lack of fidelity data. Although a process was designed for collecting this data, and the service provider did collect it, an error in data storage meant that it was not possible for the research team to see these results. This was unfortunate, as it would have been useful to consider whether the contents of the intervention as set out in the practitioner handbooks was covered with fidelity. However, the service provider has strong quality assurance procedures in place, so it is likely that fidelity was good. It would be important to amend this process in future to ensure that fidelity data was well collected and stored, so that the research team could formally monitor fidelity.

Next steps

The results of the study provide clear support for further research into the effectiveness of Talking Together. This feasibility and pilot study is one of the first steps on the path to an effectiveness evaluation. The next step should be an efficacy evaluation, using an RCT or quasi-experimental design, addressing some of the issues raised by the feasibility study and looking for evidence of impact on our primary and secondary outcome measures.

In addition, we need to explore the replicability of the programme beyond the specific areas of Bradford where it has been running to date. To explore whether the programme could be replicated in other parts of England, we are carrying out a survey of commissioning bodies regarding the need for such a programme in their area as well as barriers and facilitators to implementing a home visiting programme of this kind. We are also planning a scoping exercise to synthesise what we already know about the level and type of language provision available across different areas of England.

Conclusion

Referral rates into Talking Together indicate that it is a much-needed service, and the results of this study indicate that is positively received by the community it serves. Interpretation of our results against progression criteria would suggest that a full trial of the programme might be feasible with some adaptations particularly relating to reducing attrition. Further work exploring replicability and efficacy will move the programme further towards such a trial.

Acknowledgements

We would like to thank all the families and Language Development Workers who took part in this research. We would also like to thank Amirah Akthar, Lauren Burke, Sameera Ali, Nimarta Dharni and Rukshana Rasheed for their support. Finally, we would like to thank the Nuffield Foundation for funding the research.

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, the Ada Lovelace Institute and the Nuffield Family Justice Observatory. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org

References

1. Bercow, J. (2018). Bercow: Ten Years On: An independent review of provision for children and young people with speech, language and communication needs in England. *London: ICAN. https://www.bercow10yearson.com/*
2. Communication Trust. (2013). A generation adrift: The case for speech, language and communication to take a central role in schools' policy and practice. *London: The Communication Trust https://www.thecommunicationtrust.org.uk/media/31961/tct_genadrift.pdf*
3. Law, J., Charlton, J., & Asmussen, K. (2017). *Language as a child wellbeing indicator. London: Early Intervention Foundation.* Retrieved from http://www.eif.org.uk/wp-content/uploads/2017/09/language-child-wellbeing-indicator_Sep2017.pdf
4. Law, J., Todd, L., Clark, J., Mroz, M., & Carr, J. (2013). *Early language delays in the UK*. London: Save the Children. https://resourcecentre.savethechildren.net/node/13665/pdf/early_language_delays.pdf
5. Public Health England. (2020). Best start in speech, language and communication: Guidance to support local commissioners and service leads. *London: PHE* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/931310/BSSLC_Guidance.pdf
6. Public Health England. (2020). Early language identification measure and intervention Guidance handbook. *London: PHE*

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/939872/ELIM_Handbook_December-2020.pdf
7. Botting, N., & Conti-Ramsden, G. (2008). The role of language, social cognition, and social skill in the functional social outcomes of young adolescents with and without a history of SLI. *British Journal of Developmental Psychology*, *26*(2), 281-300.
8. Payne, A. C., Whitehurst, G. J., & Angell, A. L. (1994). The role of home literacy environment in the development of language ability in preschool children from low-income families. *Early Childhood Research Quarterly*, *9*(3-4), 427-440.
9. Dickerson J, Kelly B, Lockyer B, Bridges, S., Cartwright, C., Willan, K., Shire, K., Crossley, K., Sheldon, T.A., Lawlor, D.A., Wright, J., McEachen, R.R.C., & Pickett, K.E. (2020). Experiences of lockdown during the Covid-19 pandemic: descriptive findings from a survey of families in the Born in Bradford study [version 1; peer review: 1 approved, 1 approved with reservations]. *Wellcome Open Res*, *5*:228 (https://doi.org/10.12688/wellcomeopenres.16317.1)
10. Marmot, M., Allen, J., Goldblatt, P., Herd, E., & Morrison, J. (2020). Build back fairer: The Covid-19 Marmot review. London: Institute of Health Equity. http://www.instituteofhealthequity.org/resources-reports/build-back-fairer-the-covid-19-marmot-review/build-back-fairer-the-covid-19-marmot-review-full-report.pdf
11. Andrew, A., Cattan, S., Costa-Dias, M., Farquharson, C., Kraftman, L., Krutikova, S., Phimister, A. and Sevilla, A. (2020), 'Family time use and home learning during the Covid-19 lockdown', available online at https://ifs.org.uk/uploads/R178-Family-time-useand-home-learning-during-the-COVID-19-lockdown-1.pdf.
12. Children's Commissioner Office (2020), 'Childhood in the time of Covid', available online at https://www.childrenscommissioner. gov.uk/wp-content/uploads/2020/09/cco-childhood-in-the-time-of-covid.pdfhttps://www.childrenscommissioner.gov.uk/wp-content/uploads/2020/09/cco-childhood-in-the-time-of-covid.pdf
13. OFSTED (2020), COVID-19 series: briefing on schools, November 2020, available online at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943732/COVID-19_series_briefing_on_schools__November_2020.pdf
14. Bowyer-Crane, C., Bonetti, S., Dixon, M., Dysart, E., Newton, R., Tracey., L., Wadsworth, V. (2020, Dec 10). *Research Briefing - Early Years Settings and the COVID-19 Pandemic* https://www.niesr.ac.uk/publications/research-briefing-%C2%A0early-years-settings-and-covid-19-pandemic
15. OFSTED (2020), COVID-19 series: briefing on early years, November 2020, available online at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943492/COVID-19_series_briefing_on_early_years_November_2020.pdf
16. Axford, N., Sonthalia, S., Wrigley, Z., Goodwin, A., Ohlson, C., Bjornstad, G., Barlow., J., …& Toft, A. (2015). *Foundations for Life: What Works to Support Parent Child Interactions in the Early Years.* London; Early Intervention Foundation. http://www.eif.org.uk/publication/the-best-start-at-home.
17. Dickerson, J., Bird, P. K., McEachan, R. R., Pickett, K. E., Waiblinger, D., Uphoff, E., ... & Sahota, P. (2016). Born in Bradford's Better Start: an experimental birth cohort study to evaluate the impact of early life interventions. *BMC Public Health*, *16*(1), 711.

18. Bowyer-Crane, C., Nielsen, D., Bryant, M., Dharni, N., Heald, R., Storr, C., & Dickerson, J. (2019). A randomised controlled feasibility trial and qualitative evaluation of an early years language development intervention: study protocol of the 'outcomes of Talking Together evaluation and results'(oTTer) project. *Pilot and feasibility studies*, *5*(1), 1-12. https://doi.org/10.1186/s40814-019-0498-2
19. Local authority data received through personal communication (January 2021)
20. Hamilton, A., Plunkett, K., & Schafer, G., (2000). Infant vocabulary development assessed with a British Communicative Development Inventory: Lower scores in the UK than the USA. *Journal of Child Language*, *27*, 689-705.
21. Lancaster G, Dodd S, Williamson P. (2004) Design and analysis of pilot studies: recommendations for good practice. *J Eval Clin Pract.10*:307–12.
22. Hertzog MA. (2008). Considerations in determining sample size for pilot studies. *Res Nurs Health.;31*(2):180–91. https://doi.org/10.1002/nur.20247.
23. Sandwell Primary Care Trust. (2012). Wellcomm: A Speech and Language Toolkit for the Early Years. UK: Granada/NFER Nelson
24. Simkiss, D. E., MacCallum, F., Fan, E. E., Oates, J. M., Kimani, P. K., & Stewart-Brown, S. (2013). Validation of the mothers object relations scales in 2–4 year old children and comparison with the child–parent relationship scale. *Health and quality of life outcomes*, *11*(1), 49
25. Melhuish, E. C. (2010). *Impact of the home learning environment on child cognitive development:secondary analysis of data*. Growing Up in Scotland. Edinburgh: Scottish Government
26. Goodman R. (1997) The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry*, 38, 581–586.
27. Dale, P.S., Price, T.S., Bishop, D.V.M. & Plomin, R. (2003) Outcomes of early language delay: I. Predicting persistent and transient delay at 3 and 4 years. *J Speech Lang Hear Res*, *46*, 544–560.
28. Barnett, M. A., Gustafsson, H., Deng, M., Mills-Koonce, W. R., & Cox, M. (2012). Bidirectional Associations Among Sensitive Parenting, Language Development, and Social Competence. *Infant and Child Development*, *21*, 374–393.
29. Melhuish, E., & Gardiner, J. (2020). *Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to age five years*. Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/867140/SEED_AGE_5_REPORT_FEB.pdf
30. Girard, L. C., Doyle, O., & Tremblay, R. E. (2017). Maternal warmth and toddler development: support for transactional models in disadvantaged families. *European Child & Adolescent Psychiatry*, *26*, 497–507.