EVALUATION OF EARLY TALK BOOST EFFICACY TRIAL

EVALUATION PROTOCOL

Evaluator: Institute for Employment Studies

PI – Seemanti Ghosh

Evaluation of Early Talk Boost Efficacy Trial
Evaluation Protocol
Evaluator (institution): IES
Principal investigator(s): Seemanti Ghosh

Template last updated: December 2022

Education Endowment Foundation

Evaluation summary

Project title	Evaluation of Early Talk Boost Efficacy Trial
Developer *(Institution)*	Speech and Language UK
Evaluator *(Institution)*	Institute for Employment Studies
Principal investigator(s)	Seemanti Ghosh
Protocol author(s)	Susie Bamford, Allie Nancarrow, Clare Huxley, Jade Talbot, and Seemanti Ghosh
Trial design	Two-arm cluster randomised waitlist-controlled trial with random allocation at setting level
Trial type	Efficacy
Pupil age range and Key stage	3-4, Key Stage – Early Years
Number of schools *(at design stage)*	132
Number of pupils *(at design stage)*	660
Primary outcome measure and source	Expressive Communication Subscale score (Preschool Language Scale 5)
Secondary outcome measure and source	Outcomes of Communication Under Six Short Form score (FOCUS-34)

Protocol version history

Version	Date	Reason for revision
1.0 [*original*]	August 2023	N/A

Contents

Tables

Figures

Appendix tables

Study rationale and background

Language and communication skills tend to impact children's academic achievement and later employment (Law, Charlton and Asmussen, 2017). In the UK, approximately 85,000 to 90,000 2 to 6-year-olds are referred to speech and language therapists annually Children from low socioeconomic backgrounds are more than twice as likely than their more advantaged peers to have speech, language and communication needs (Law, Charlton and Asmussen, 2017). The higher prevalence of language difficulties in children from low socioeconomic backgrounds is most likely a component of the achievement gap that is present in the Early Years and continues to exist throughout the school years. However, speech and language interventions are often put in place once children have been referred for speech and language therapy. The children who have mild language development issues are notably underserviced (Marshall et al., 2017) and there are not enough interventions targeted at this group. Additionally, COVID-19 has contributed to the exacerbation of the speech and language development gap among children from deprived socioeconomic backgrounds (González & Bonal, 2021).) Therefore, more than ever there is a need for interventions that cater to children who are facing some issues in language development but not enough to be sent to therapy to support them with their language development. Early Talk Boost is a targeted intervention aimed at 3–4-year-old children who are behind in their language and communication, boosting their language skills to help narrow the gap between them and their peers. The intervention has been designed by specialist speech and language therapists and a highly experienced nursery teacher and is based on strong evidence of what is known to support language development in the Early Years. The programme is delivered by Speech and Language UK, previously known as I CAN.

A small randomised controlled trial (RCT) of Early Talk Boost was carried out between 2015 and 2017, including 15 settings and 85 children (Reeves et al., 2018). Early Years settings were randomly assigned to use the programme or to act as waitlist comparison group (who received the intervention in the following term). Children were assessed before and after the completion of the programme, using the Preschool Language Scales, a standardised assessment of expressive and auditory (receptive) language skills. There was a significant progress made within the treatment group over time, showing 6 month's additional progress(Speech and Language UK, 2015). When rigorously assessing the impact of exposure to the programme, by comparing the impact between treatment and control groups, the effect size of 0.145 would equate to 2 additional months progress for pupils in the treatment group (Reeves et al., 2018). Children in the treatment group made more progress in both expressive and receptive language than pupils in the control group. The Early Intervention Foundation's Guidebook gives this study a Level 2 evidence rating, noting that the programme has evidence of improving a child's outcome through a study involving at least 20 participants, representing 60% of the sample, and using validated instruments. The programme does not receive a Level 3 evidence rating as the RCT was not assessed by EIF as having been rigorously conducted.

Wider evidence in the Early Years' literature provides support for elements of the logic model, including evidence that approaches that develop children's vocabulary and ability to develop conversations can support early language outcomes (Law, Charlton and Asmussen, 2017). The EEF's Early Years Toolkit notes that communication and language approaches that emphasise the importance of spoken language and verbal interaction for young children can

have very high impacts, equivalent to 7 months' additional progress, on average, based on extensive evidence.

Given the research and practice gap for a language programme for 3-4-year-old children who are behind in their language development, EEF are funding an efficacy trial of the programme to provide further evidence for what works for this target group and to support education recovery following the pandemic. This evaluation aims to learn from the findings of 2015 evaluation and improve on its limitations. For example, the 2015 evaluation selected children into the study firstly based on practitioner nomination, followed by selection based on a form filled in by practitioners based on their observation of the children. This could have invited biases that cannot be accounted for. This evaluation will be using a screening instrument that is specifically meant to identify children with language development issues and this will be administered by an independent third party. Furthermore, communication will also be measured using a standardized scale in this trial vis-à-vis practitioner observations. This evaluation has been designed as an efficacy trial, with a much larger sample size and sufficiently powered to detect an appropriate minimum detectable effect size which is an improvement over the 2015 small-scale trial. Therefore, the results are expected to be robust.

The aim of this evaluation will be to assess, using a Randomized Controlled Trial, if Early Talk Boost is able to improve communication and language outcomes among 3–4-year-old children who are behind.in terms of language development, in particular in their talking and understanding words. In addition to this, a thorough implementation and process evaluation will aim to establish fidelity in implementation, explore key stakeholder engagement with the programme, identify any factors influencing implementation that may facilitate/hinder impact at the child-level (including wider national/policy contexts), and to inform and contextualise the quantitative findings. IPE data will be used to explore and interrogate aspects of the Theory of Change such as evidence regarding change mechanisms, perceived short-term outcomes and other assumptions. In line with the Theory of Change, we will examine the experiences and behaviour of the three key stakeholders: practitioners, children, and parents, as well as interviewing the delivery team.

Intervention

Intervention Name

Early Talk Boost

Why (Theory and rationale)

Early Talk Boost is an intervention that works in partnership with Early Years settings to provide Early Years practitioners with specialised training in a speech and language intervention targeted at 3–4-year-old children. The programme is designed and delivered by Speech and Language UK. As outlined in the previous section, there is evidence that children who participate in Early Talk Boost gain positive impacts in their expressive language and communication skills. Being able to talk and understand words are key skills for young children enabling them to access all areas of the Early Years Foundation Stage Curriculum (EYFS). A proportion of children struggle with these skills and are behind compared to their peers. These children are also more likely to come from disadvantaged backgrounds, and the recent COVID pandemic has only exaggerated this gap (González & Bonal, 2021). With the Early Talk Boost intervention, Speech and Language UK aim to boost the language skills of 3–4-year-old

children who are behind in their language and communication, helping them to catch up with their peers.

Who (Recipients)

This intervention is targeted at 3–4-year-old children who are behind their peers in their language and communication development. Typically, these children are identified by early years practitioners who have undergone training with Speech and Language UK - so that they use specialised skills and informed judgement to select the children that would most benefit and then they go on to monitor their progress with a tailored tracker. This training is an important part of the intervention. However, practitioner observation can invite biases that will impede the process of a robust evaluation of the program. As this is an efficacy trial, only half of the settings and children will receive the intervention in the first instance, it was important to identify a consistent mechanism that minimises bias in identifying eligible children. The WellComm Early Years Speech and Language Toolkit will be used as a screening tool for all children in the target age range at each setting. This screening tool has been selected, after giving fair consideration to multiple early years language measures such as the Oxford Language Screen, because it is age-appropriate and can be used to identify children who are experiencing early speech and language difficulties

In this trial we are aiming to recruit 132 participating settings across two geographical regions covered by Stronger Practice Hubs in the East of England and South West. All 3- to 4-year-old children in the 132 settings will be invited to participate in the screening test and at this point parents will be provided with clear information about the project and have the right to withdraw the consent. Children will then undergo the WellComm screening test and only children who meet the cut off (as described in the exclusion criteria section) will be selected into the study. More details on the selection process is provided below in the Impact Evaluation Design section. Once the screening has been completed and children have been selected into the study, a baseline measure of the primary outcome (expressive language) will be conducted for all selected children. Following which, 66 settings will be randomised to receive the intervention and 66 settings will be randomised to be the control group. Therefore, as part of the evaluation, only the selected children in the 66 intervention settings will receive Early Talk Boost.

Usually, Speech and Language UK aim to deliver to a maximum of 8 children per setting. However, for the purpose of the evaluation, we have powered the study to detect the Minimum Detectable Effect Size (MDES) with an average of 5 children per setting as that is the anticipated average number of children who might qualify for the intervention. The upper limit of group size is 8 children as that is the maximum size the intervention has been designed for. The lower limit of the group size has been set as 3, as a group size below that would render the intervention delivery to be costly.

What (materials)

Intervention settings will receive an Early Talk Boost Intervention Pack. The pack includes most of the items and toys that settings need to run the intervention. Specifically, the pack includes:

- The Early Talk Boost Intervention Manual containing detailed group session plans and required resources
- Books - 10 Jake & Tizzy Complete Sets (Titles 1-8) that have been specifically designed and written for the intervention
- Early Talk Boost Participant Book, which accompanies the training
- Toolkit Bag including most of the resources needed for intervention activities
- Access to the Online Tracker where progress can be monitored

What (procedures)

The Early Talk Boost programme starts with training for Early Years practitioners, which includes four main topics. These are:

- typical speech, language and communication development;
- things that can potentially go wrong and what impact that has on children's learning and behaviour;
- which children will receive the most benefit from the intervention and how to use the Tracker to identify them and track their progress;
- planning and running the Early Talk Boost sessions, along with using the manual and accompanying resources, as well as involving parents (i.e. parent workshops and using the Jake and Tizzy book to support the home learning environment).

The training develops practitioners (1) ability to understand the Early Talk Boost activities and how to support children's language, (2) understanding of how to identify children who will benefit from Early Talk Boost and monitor their progress, and (3) understanding of working with parents to support their child's talking and understanding words. Two practitioners per setting are invited to take part in the training, which allows for continuity in the settings in case of staff changes and supports the practitioners in establishing Early Talk Boost as part of the daily nursery routine and sharing of good practice. Only one practitioner is required to run the intervention sessions.

The practitioners then go on to deliver a nine-week intervention of three sessions per week, with a group of up to eight children in their setting. These sessions include activities and games supported by storybooks, purposely designed to boost the children's language skills, helping them catch up with their peers. Each 15-20-minute session includes two activities: a song and a story covering the key elements of language, including attention, and listening, developing vocabulary, building sentences and having conversations. Speech and Language advisors (i.e. either specialist speech and language therapists or teachers with expertise in supporting speech and language challenges) will also provide three support sessions for each setting: one half-day in-person visit and two shorter online meetings. The purpose of these is to ensure the practitioners can deliver the intervention and to support the research visits at the pace required by the study timescales. The support sessions will include the following: registering on the online tracker system and uploading children's tracker data; delivering the parent workshop; answering any queries about the intervention sessions; and capturing information on fidelity and attendance.

The intervention also engages parents/carers, with advice and support to share the books with their children at home and a parent workshop lasting 30 – 40 minutes and delivered by the trained Early Talk Boost practitioners in the settings. The parent workshop can be delivered in a couple of formats: one in which practitioners deliver scripted PowerPoint slides or one in

which parents are presented with a recorded delivery of the slide contents. Both of these formats include opportunities for parents to engage in activities. The workshop coach parents on how best to use the Jake and Tizzy books at home and about contingent language behaviours and the associated use of the books and go on to use and are encouraged to enjoy the books with their children at home. In the short-term, parents use new targeted contingent communication with their children whilst sharing the books. However, since parental engagement was not measured the 2015 evaluation, and because the more intensive in-setting elements of the programme are considered to be the core components of the programme most likely to be driving impacts on children's outcomes, this study has not been designed to capture the impact of parental engagement on the outcomes,

Early Talk Boost is accompanied by a tracker based on developmental norms (which is usually used by the practitioner to select children for the intervention but will not be used for selection in this trial) and as a post-intervention progress measure. The tracker scores can be uploaded to an online tool that analyses the results at an individual and cohort level.

More information about Early Talk Boost can be found on the Speech and Language UK website: https://speechandlanguage.org.uk/training-licensing/programmes-for-nurseries-and-schools/early-talk-boost/

Who (providers)

Early Years setting staff who have participated in Early Talk Boost training provided by Speech and Language UK will deliver the intervention in settings. The Early Talk Boost training is delivered in two online sessions. Two Early Years staff from each participating setting must take part in the training.

How? (mode of delivery)

The Early Talk Boost intervention is delivered in small groups (maximum of 8) of targeted children.

Participating children receive the intervention from their Early Years setting staff during their usual attendance at their Early Years settings. It is also hoped that they receive parental input through looking at and talking about the Jake and Tizzy books at home.

Where (Location)

Children receive the intervention in Early Years settings. Additionally, parents/carers are encouraged to support their children's communication and language development at home.

When and how much (duration and dosage)

The intervention lasts 9 weeks during term time and consists of three 15-20 minute targeted sessions per week. Books are sent home for parent-child storytimes. Ideally the intervention is delivered over 9 consecutive weeks but a maximum 2-week gap maybe included for example to accommodate school holidays.

Tailoring and adaptation

The manual allows for adaptation if the tasks are too hard or too easy for individual children taking part. These are included in the text under 'step up' and 'step down' sections No additional tailoring or adaptation is planned.

How well (planned)

- Typically post training support is provided by trainers who are licensed by Speech and Language UK to deliver the Early Talk Boost training to settings or by the commissioning Local Authority. In this study additional support will be provided by a Speech and Language UK Advisor comprising three support sessions, the first two of which will be online and the third may be online or in-person depending on logistics.
- Settings on the waiting list for the intervention will be involved in regular updates and a keeping in touch programme.
- Speech and Language UK will set up additional support and link settings together through a dedicated social media channel.

Logic Model

The logic model was developed during the trial set-up phase and is displayed below in Figure 1, with a target population defined as typically developing 3-4-year-old children living in disadvantaged areas who demonstrate difficulties with communication and language. Details regarding assumed causal mechanisms are provided in Appendix A.

The logic model describes the intervention activities and illustrates the causal mechanisms underpinning the intervention as well as the anticipated short-term, intermediate and longer-term outcomes. The model suggests that the intervention changes and improves practitioner and parent behaviour around interactions and language practices and that because of these changes, children will have improved attainment in communication and language.

The logic model shows how practitioners will receive training and resources for Early Talk Boost and how they go on to deliver this to children in their setting. This leads to a change in the way practitioners engage with children in the short term by having a greater understanding of SLC development and how to support it, modelling language and interaction skills, using strategies to improve language, and tracking children's progress. Longer term outcomes will continue as a greater understanding of the importance of language and communication and how to spot when a child is falling behind, and a subsequent change in behaviour to address this.

For children, the inputs are taking part and enjoying the targeted Early Talk Boost sessions, being exposed to metalinguistic opportunities, and experiencing the Jake and Tizzy books and associated activities in the setting and at home. This allows children to experience models of listening and communication with positive feedback, and to practice these new skills. This leads to improved attention and listening, improved social interaction behaviours, and improved learning and engagement in their setting. Longer term outcomes include accelerated progress and attainment in language and communication, children are more ready to learn when they start school, and barriers to learning and making friends are reduced.

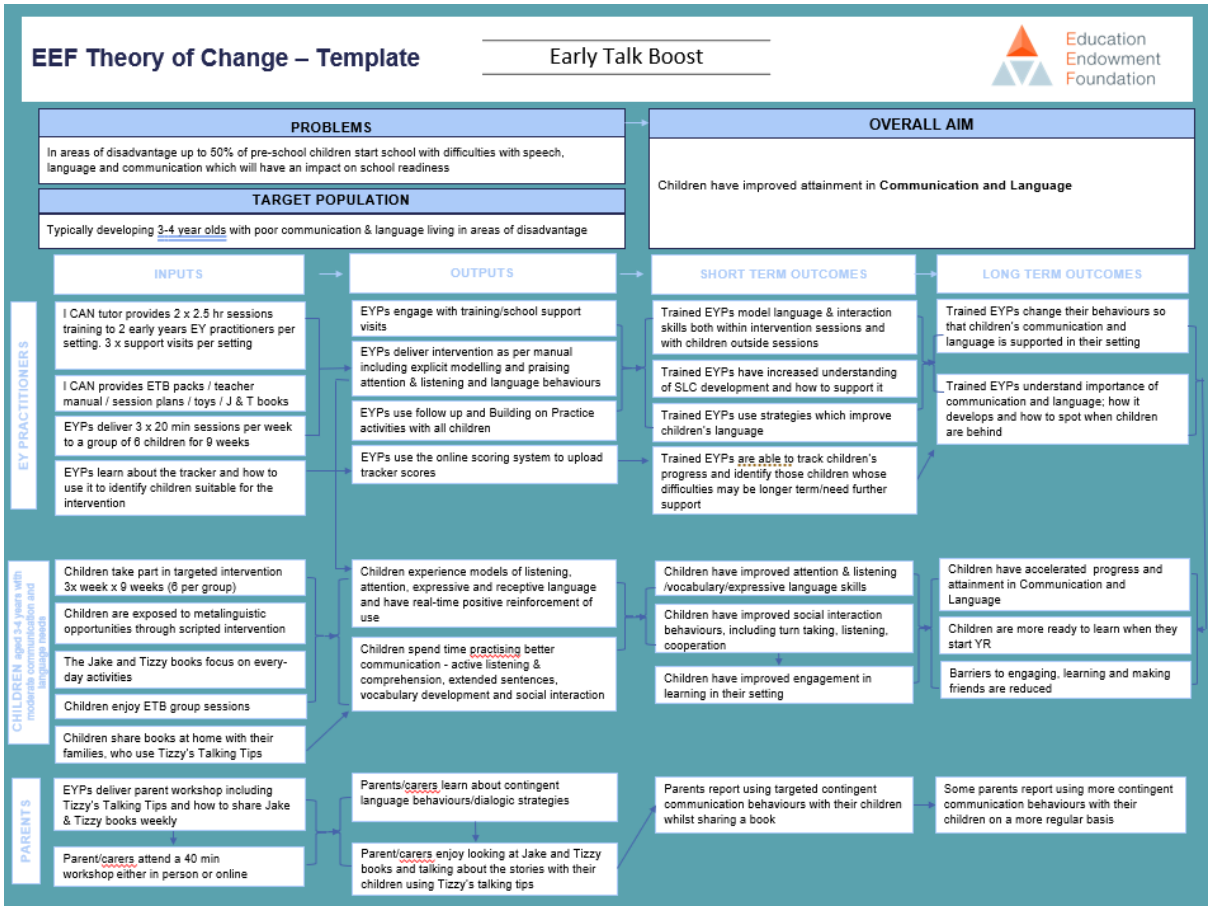


Figure 1: Logic model

Impact evaluation design

Primary research questions

RQ 1) What is the difference in expressive language ability measured by the Preschool Language Scale 5 Expressive Communication subscale, between children in intervention settings receiving Early Talk Boost and children in waitlist control settings receiving business as usual, after controlling for the baseline measure of outcome?

RQ 2) What is the difference in attention and social communication measured by the FOCUS-34, between children in intervention settings receiving Early Talk Boost and children in waitlist control settings receiving business as usual?

Secondary research questions

The secondary research questions will explore heterogenous effects by various sub-groups:

RQ3) Early Years Pupil Premium (EYPP) eligibility: EYPP, non-EYPP

a. Does the impact of the Early Talk Boost programme on expressive language vary by the EYPP eligibility status of the children?
b. Does the impact of the Early Talk Boost programme on attention and social communication vary by the EYPP eligibility status of the children?
c. Does Early Talk Boost programme have a significant impact on the expressive language of EYPP children? (We have not powered the study for this, but we will explore)
d. Does Early Talk Boost programme have a significant impact on the attention and social communication of EYPP children? (We have not powered the study for this, but we will explore)

RQ4) Setting type: Private, voluntary, and independent (PVI); Maintained

a. Does the impact of the Early Talk Boost programme on expressive language vary by the setting type (PVI vs Maintained)?
b. Does the impact of the Early Talk Boost programme on attention and social communication vary by the setting type (PVI vs Maintained)?

RQ5) Quartiles of baseline measure of outcome:

a. Does the impact of the Early Talk Boost programme on expressive language vary by the quartile of the baseline measure of the outcome?
b. Does the impact of the Early Talk Boost programme on attention and social communication vary by the quartile of the baseline measure of the outcome?

RQ6) English as an additional language (EAL):

a. Does the impact of the Early Talk Boost programme on expressive language vary by EAL status?
b. Does the impact of the Early Talk Boost programme on attention and social communication vary by EAL status?

RQ7) If we are able to collect information on practitioners receiving Department for Education (DfE) training, such as those listed here[1], we will explore:

a. Does the impact of the Early Talk Boost programme on expressive language vary by practitioners having received DfE training?
b. Does the impact of the Early Talk Boost programme attention and social communication vary by practitioners having received DfE training?

Design

Table 1: Trial design

Trial design, including number of arms		Two-arm, cluster randomised
Unit of randomisation		Setting
Stratification variables (if applicable)		Stronger Practice Hub
Primary outcome	**Variable**	Expressive language
	Measure (instrument, scale, source)	Preschool Language Scale 5 - Expressive Communication Subscale (standardised score)
Secondary outcome(s)	**Variable(s)**	Attention and social communication
	Measure(s) (instrument, scale, source)	FOCUS-34, (7 point Likert scales, 34 items, 0-238) (standardised score)
Baseline for primary outcome	**Variable**	Expressive language
	Measure (instrument, scale, source)	Preschool Language Scale 5 – Expressive Communication Subscale (raw score, standardised score)
Baseline for secondary outcome	**Variable**	No baseline – post-test only
	Measure (instrument, scale, source)	No baseline – post-test only

[1] As part of the early years education recovery package, DfE supports 5 offers, which focus on supporting with continuing professional development:

- early years child development training
- professional development programme phase 3 (PDP3)
- Nuffield Early Language Intervention (NELI)
- national professional qualification in early years leadership (NPQEYL)
- home learning environment (HLE)

The evaluation will be conducted as a two-arm cluster (setting-level) randomised waitlist-controlled trial to identify the impact of the Early Talk Boost programme on expressive language, and the attention and communication of 3-4-year-old children who are experiencing delay in speech and language development. Given this is an efficacy trial, an effort has been made to maximise the likelihood of observing an intervention effect if one exists.

The trial will be conducted in nursery settings under two Stronger Practice Hubs - the East of England Early Years Stronger Practice Hubs and Early Years South West Stronger Practice Hub. Given the nature of the intervention, we believe that a setting-level randomisation is best suited for the evaluation design rather than an individual-level randomisation. This allows us to prevent contamination and also to ease delivery in Early Years settings. In conducting the randomisation, to minimise the risk of imbalance across treatment and control groups, we recommend stratification by Stronger Practice Hubs. Randomisation will occur after baseline assessments have been completed to minimize potential attrition bias in ITT analysis.

Settings randomized into the intervention group will receive the Early Talk Boost program that will run for 27 sessions or 9 weeks (3 sessions per week; minimum of 7 weeks) and the settings randomized to the waitlist control group will be business-as-usual for the duration of the trial. Following completion of the trial, the settings in the waitlist control group will have the opportunity to receive Early Talk Boost in the following academic year.

Early Talk Boost was designed to enhance children's language, attention and communication skills. The primary outcome of interest is expressive language as measured by the Expressive Communication subscale of the PLS-5 and the secondary outcome is attention and communication as measured by the FOCUS-34. These measures have been carefully chosen after in depth discussion with the delivery team to align with the logic model and the perceived and intended effects of the intervention.

Participant selection

The intervention is designed for 3–4-year-old children in Early Years settings who are behind in their talking and understanding of words.

Setting Level

Eligibility:

Settings in the following Local Authorities (LAs) can take part in the study: Norfolk, Suffolk, Bournemouth, Essex, Southend-on-sea, Cambridgeshire, Peterborough Christchurch and Poole, Cornwall, Devon, Dorset, Plymouth, and Torbay.

Both Private, Voluntary and Independent (PVI) and maintained nurseries (either stand alone or as part of a Primary School) are eligible.

Settings can only sign up to receive one funded Stronger Practice Hub programme between 2022-2025.

Inclusion criteria:

Settings need to be able to:

- Commit 2 practitioners to attend the training

- Deliver the intervention over 3 sessions per week for 9 weeks
- Have at least 8 children who are between 3 and 4 years old (date of birth must fall between 1 December 2019 and 1 October 2020)
- Commit to supporting evaluation activities.

Exclusion criteria:

Settings can't take part if they:

- Are already running Early Talk Boost groups
- Are taking part in any other trial funded by the EEF or similar funder or are taking part in another SPH-funded programme.

Child Level

Inclusion criteria:

Children need to be:

- 3-4 years old who may be behind with talking and understanding words.
- Has to be signed up for at least 2 terms.
- Should be able to attend the sessions at least 3 times a week.
- 3 years old and no more than 4.5 years old at the end of the intervention. Only children with birth date between 1 December 2019 and 1 Oct 2020 will be included e.g., the oldest children will just turn 4 on 1 Dec 2023 (and won't be 4.5 until end May of that year so will have finished the intervention by then). And the youngest children will be born on 1 October 2020 and will just turn 3 on 1 Oct 2023 (allows for baselines from 1 Oct 2023).

Exclusion criteria:

Children can't take part if they have:

- SEND concerns - have significant SEND which would prevent them from accessing the assessment and/or would be distressed through completing the assessment and taking part in the intervention.
- Diagnosed conditions e.g., Down's Syndrome, complex learning needs.

To eliminate practitioner biases in terms of screening children eligible for the Early Talk Boost intervention, the WellComm Early Years Speech and Language Toolkit will be used as a screening tool for all children at each setting. All children in 132 participating settings will be invited to complete the WellComm assessment , administered by Qa as a screening tool which will determine their current language level. At this point parents will have the option to withdraw consent.

The WellComm was designed for use with children aged 6 months to 6 years. This measure takes approximately 20 minutes to complete and includes a total of nine sections, with each section containing a total of 10 items. Each section has possible scores of 0 to 10. For each section, scores are distributed across three RAG ratings: green (i.e. no intervention needed), amber (i.e. extra support needed) and red (i.e. suggested SLT referral). For example, if a child receives a score of 8 to 10 in section 1, they receive a green rating. A score of 6 to 7 in section

1 grants children an amber rating, and a score of 0 to 5 in section 1 is a red rating. Age determines where in the measure children start, and the administrator moves back sections until a green rating is achieved. Children given amber and the top end of red ratings for their age will be eligible for the Early Talk Boost programme. The maximum eligible group size possible is 8, therefore, we will try to recruit 8 children per setting using this eligibility criteria. If more than eight children are identified as eligible, eight children will be randomly selected to be tested at baseline, with an aim to recruit at least 3 children with EYPP. If less than 3 children are identified in a setting through this process, that setting will be dropped from the study. However, on average we anticipate 5 children per setting to fall into the eligible category, using the WellComm screening tool acting as the eligibility criteria.

The WellComm has been norm-referenced and is currently widely used by practitioners. Its concurrent reliability with the RDLS III is .89, and it correlates with the PLS-4 auditory subscale. According to GL Assessment, from a sample of a total of 83,000 three- to four-year-olds in the last five years, 46% of children receive a green rating, 20% receive an amber rating and 34% receive a red rating. The overarching aim of the screening process will be to identify children who score slightly below expectations for their age will go on to participate in the trial.

The Delivery team, Speech and Language UK will be leading the recruitment supported by the Evaluation team. Flyers advertising the study, alongside setting level information sheets will be sent out to settings in the two Stronger Practice Hub regions. Interested settings reply with an expression of interest to Speech and Language UK. Speech and Language UK. Settings are then invited to an information session, where they find out more about Early Talk Boost ad the evaluation. Following this session, they are sent the Memorandum of Understanding (MOU) forms to read and sign. Speech and Language UK keeps and updates a database of signed up settings and Speech and Language UK transfers the database of settings to IES. IES send participating settings parent information sheets to distribute to their parents allowing the parents the opportunity to read about the study and opt their children out if they wish. Settings return child level data of participating children to IES. IES then has a database of participating settings and children. This can then be used to arrange for screening assessments to determine the final list of participating eligible children.

Screening table

Measurement instrument	Timepoint(s)	Collected by
WellComm	October – November 2023	Qa

Incentives

Nurseries taking part in the programme will receive an Early Talk Boost Intervention pack which normally costs £550 and training and support, all of which is fully funded. Control settings will receive the same fully funded training and support in the academic year 2024/2025. In recognition that there are costs associated with undertaking and implementing professional development, additional funding has been made available by the EEF to reimburse nurseries for approximately 50% their staff's time/cover cost to attend training, by providing settings with £60 per practitioner who completes the professional development activities, for up to two staff members. In addition, settings will receive £400 in the evaluation year (2023/2024) to assist in parental recruitment and completion of evaluation measures

(surveys, setting visits, child language assessments). This will be divided into £150 for baseline (i.e., Spring 2024) and £250 on completion of endline (by Aug/Sept 2024).

Outcome measures

Baseline measures

The Preschool Language Scale 5 UK edition (PLS-5 UK)[1] is suitable for children from birth up to seven years and eleven months' old. As the intervention will be delivered with 3-4-year-old children, this should allow a sensitive measure of language change with low risk of ceiling or floor effects. It includes two standardised scales (Auditory Comprehension and Expressive Communication), and three supplemental measures (Language Sample Checklist, Articulation Screener, Home Communication Questionnaire). The test is administered by speech and language therapists. The PLS-5 UK was published in 2014, with data collected from a representative normative sample in the UK from ages 2.6 years to 7.7 years; UK norms were extrapolated for ages 7.8 to 7.11 years so items should reflect current language use. A previous, smaller-scale evaluation of the Early Talk Boost intervention (Reeves et al., 2018) used the previous version of the Preschool Language Scale (PLS-4; Zimmerman et al., 2009) as the primary outcome measure and found this to be sensitive for measuring relevant changes in language and communication.

In order to reduce the testing burden on participating children and settings, we will use the Expressive Communication subscale from the PLS-5 as the baseline measure for the primary outcome. Expressive language is an outcome identified in the Theory of Change and has been identified by the delivery team as a key area where they would expect to see impact from the Early Talk Boost intervention. The Expressive Communication subscale contains 67 items that assess children's vocal development and social communication (Hsiao et al., 2021). Preschool-age children and children in Early Years education are asked to name common objects, use concepts that describe objects and express quantity, and use specific prepositions, grammatical markers, and sentence structures. Age determines where children start (in six-month blocks; e.g., 3 years – 3 years 5 months; 3 years 6 months – four years). If children struggle with the items, then the administrator will go back until a baseline is found. The PLS-5 has been norm referenced, and its concurrent validity with the PLS-4 is .80 for the Auditory Comprehension and Expressive Communication subscales. The correlations between the PLS-5 and the CELF ranged from .70 to .82.

The baseline test will be delivered by speech and language therapists from ELKAN. As the baseline assessments occur before randomisation, assessors will be blind to eventual trial condition. Raw data will be provided by Elklan, and the evaluation team will compute scores. We are unable to provide the instrument because it is a commercial test and not publicly available.

Primary outcome

The Expressive Communication subscale from the PLS-5 will be administered at pre- and post-test as a primary outcome measure. As described in the logic model, the intervention is expected to raise children's expressive language skills, and the primary outcome measure will address the first research question.

Speech and language therapists at Elklan would administer the PLS-5 assessments at both baseline and endline. Assessors would be blind to treatment/control condition at both time points.

Secondary outcomes

The secondary measure will be attention and social communication and we will measure these using the Focus on the Outcomes of Communication Under Six (FOCUS-34) short form (34 items)[2]. The Focus on the Outcomes of Communication Under Six (FOCUS) was designed for use with children aged 1 year and 6 months to 6 years. The FOCUS assessment is available as a 50 item and 34 item versions. The 34-item version retains the validity of the fuller 50 item measure. In order to reduce burden on children and settings, we will use the FOCUS-34 as this takes less time to administer. The short version contains 34 items and takes approximately 7 to 15 minutes to complete. The correlation between the short version and the longer, original (50-item) FOCUS was demonstrated to be .98 (Oddson et al., 2019).

We will administer this at post-test only in order to minimise the testing burden on children and settings. The FOCUS-34 is suitable for use with children aged one year and six months up to six years so should capture a wide range of scores for 3-4-year-old children. The FOCUS-34 aims to evaluate communicative participation in preschool children including: expressive language, pragmatics, receptive language / attention, intelligibility, social/play, independence, and coping/emotions. The FOCUS has been normed with English-speaking children in Canada (Thomas-Stonell et al., 2010). Outcomes identified in the Theory of Change logic model included improvements in children's attention and listening, and social communication behaviours, and this measure will address the second research question.

The FOCUS-34 assessment is available as a parent or clinician/practitioner version. We will ask setting practitioners to administer the practitioner/clinician version of the test with the participating children at their setting, i.e., an average of 5 but up to a maximum of 8. Assessing wider communication skills requires knowledge of the child being assessed and as children will have been attending settings for several months by endline, practitioners should have sufficient knowledge to answer the FOCUS-34 assessments. The evaluation team will provide online FOCUS-34 forms for each child that can be filled in online and where practitioners can save and then return to a form at a later point as needed. This approach has been used to minimise burden on practitioners. It was felt that asking parents to complete the FOCUS-34 form would constitute unacceptable burden and likely yield low response rates.

The table below summarises the outcome measures, including instruments and timepoints.

Outcome measures table

Outcome	**Measurement instrument**	**Timepoint(s)**	**Collected by**
Primary	PLS-5 Expressive Communication subscale	Baseline and endline	EKLAN
Secondary	FOCUS-34	Endline	Practitioners

[1] Additional information about the PLS-5 can be found online here.

[2] Additional information about the FOCUS-34 can be found online here.

Data collection

Setting- and practitioner-level data

The delivery team, Speech and Language UK, will collect the following setting-level data during recruitment and share with the evaluation team, IES, via secure file sharing folder on the IES OneDrive: name of setting Local Authority, name of setting, setting address/postcode, URN (if applicable), setting type (PVI, School Nursery (SN), other), Ofsted status, setting email address, setting main contact name, setting main contact email address, head teacher / nursery manager name, head teacher / nursery manager email address, head teacher / nursery manager phone number, number of 3- and 4-year-olds, and whether they are currently or planning to be involved in any other Early Years interventions. Speech and Language UK will also collect the following monitoring data and share with IES via OneDrive: practitioner training attendance data, support session data (i.e. method and number), number of sessions delivered to children per setting and session attendance.

IES will collect the following setting-level data: setting level usual practice survey data (as reported by head teacher / nursery manager), practitioner usual practice survey data, practitioner highest qualification, practitioner role in setting, practitioner years of experience in early years, practitioner receipt of specific DfE training, setting staff interview / setting visit recordings, setting staff interview / setting visit transcripts, setting staff views of the programme and setting observation data. These data will be collected via surveys, monitoring data, and Zoom / Teams / Dictaphones / phones.

Child-level data

IES will collect the following child-level data from settings via email/phone: child first name, child last name, child sex, child DOB, child home postcode, child EAL status (if available), child EYPP status (if available), Unique Pupil Number (UPN; if available), Unique Establishment Number (URN) or LAESTAB (if available), name of setting child attends, child attendance (i.e. voluntary withdrawal from study and passive withdrawal by moving house / settings / leaving nursery) and parent/carer interview data (with up to 10 interviews with parents or carers). Speech and Language UK will collect child attendance data in the form of the Early Talk Boost register and will share with IES via OneDrive. Subcontractors who are administering the WellComm and PLS-5 measures will collect child data at settings and share with IES via OneDrive. Practitioners will complete the FOCUS-34 and share the data with IES via OneDrive.

Developer-level data

IES will collect developer-level data in the form of interviews with the delivery team via Zoom / Teams / phone.

Sample size

We conducted a sample size calculation to understand the Minimum Detectable Effect Size (MDES), helping us to determine the feasible number of variations we could use to detect a

main effect of at least 0.2 Standard Deviation. The sample size calculation was conducted based on Oughton (2022) and in Table 2 we present 3 variations.

In previous early year trials the setting-level intra-cluster correlation coefficient (ICC) ranged between 0.19 and 0.11 (Every Child Counts, Torgerson et al, 2022; Maths Champions, Robinson-Smith et al., 2018), so to be conservative an ICC of 0.17 will be used. As there is a lack of evidence to inform a pre-post correlation, the MDES was calculated using a range of values (between 0.5 and 0.7). We present here the calculations with pre-post correlation 0.6. Similarly, as there can be high levels of child-level attrition in EY trials, a variety of these were explored (10% to 25%), but a conservative of 23% child-level attrition has been used for the below sample calculation.

Assuming pre-post correlation of 0.6, intra-cluster correlation of 0.17 and 132 settings, we present 3 variations with assumptions of 8, 6 and 5 children per setting respectively in Overall 1, 2 and 3. As shown, the MDES that can be detected in each of these cases are 0.20, 0.22 and 0.23 respectively for Overall 1, 2 and 3 without attrition. Assuming 15% setting-level attrition, 23% child-level attrition, these MDES would move up to 0.23, 0.25 and 0.26 respectively.

The largest cluster size per setting in this evaluation can be 8 as this is the maximum group size for the intervention to run. However, 5 is the anticipated average cluster size but it is not ruled out that will be settings where the actual cluster size is either smaller or larger. Given the upper limit of 8 children, if a setting has more than 8 eligible for the intervention, 8 will be selected randomly. If there are less than 5 children in a setting, the lower limit would be 3, given the mean over 132 settings still remains 5. An average of below 5 will lead to an underpowered study and the evaluation will not be able to detect an effect if present below 0.23 (without attrition) and 0.26 (with attrition).

In Table 2 we also present sample size calculation to understand the MDES for EYPP eligible children only. We assumed 20% EYPP eligibility which is a conservative assumption, given evidence suggests 10% (https://explore-education-statistics.service.gov.uk/find-statistics/education-provision-children-under-5). Which leads to 1 child per setting, if we consider the average setting size to be 5. Assuming, 20% EYPP eligibility (total of 132 children) and 0.6 pre-post correlation, we get MDES 0.39, which is a very large MDES. Therefore, we conclude that this trial is not powered to detect an effect on the EYPP-subgroup.

In cases where settings have more than 8 children who are eligible for the intervention, we will consider oversampling the EYPP subgroup as this group is of primary interest. Where possible, we want to include at more than one EYPP child per setting to conduct EYPP subgroup analysis.

Table 2: Sample size calculations

		Overall 1	Overall 2	Overall 3	EYPP
Minimum Detectable Effect Size (MDES)[2]		0.20	0.22	0.23	0.39
	level 1 (pupil)	0.6	0.6	0.6	0.6

Pre-test/ post-test correlations	level 2 (class)				
	level 3 (setting)				
Intra-cluster correlations (ICCs)	level 2 (class)				
	level 3 (setting)	0.17	0.17	0.17	0
Alpha[3]		0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8
One-sided or two-sided?		2	2	2	2
Average cluster size		8	6	5	1
Number of schools[4,5]	Intervention	66	66	66	66
	Control	66	66	66	66
	Total	132	132	132	132
Number of pupils	Intervention	528	396	330	66
	Control	528	396	330	66
	Total	1056	792	660	132

Randomisation

Method

We propose a two-arm cluster randomised trial and Early Years setting will act as the unit of cluster. Given the nature of the intervention, we believe that a setting-level randomisation is best suited for the evaluation design rather than an individual-level randomisation to prevent contamination and also ease delivery in Early Years settings. Settings will be randomised to receive the Early Talk Boost intervention, or to be in the waitlist control group where they will continue with business as usual and then receive the intervention the following term. Randomisation will occur after baseline assessments have been completed to minimize potential attrition bias in ITT analysis.

We will conduct a stratified randomisation using the Stronger Practice Hub areas as a stratum to ensure that treatment and control groups are balanced across areas. As this is a cluster-randomised trial, settings (as clusters), will be randomly allocated 1:1 to receive either the intervention or business as usual.

Randomisation will be conducted by IES statisticians using code in standard software such as STATA/R which will be stored for reproducibility and transparency and published as part of the SAP and the final report. The statistician will not be blinded to treatment group. They will pass this information to the delivery team so that they can commence training and delivery.

Establishing whether randomisation has been successful

We will assess for imbalance on the following variables:

- Age, gender and ethnicity of the pupil
- Baseline measure of primary outcome collected using PLS-5

- EY PP eligibility
- PVI and Maintained categories

Statistical analysis

Primary analysis

We propose estimating primary and secondary outcomes using multi-level modelling, to take into account the nested structure of the data. The primary outcome will be measured at child level and the analysis will control for baseline values of the outcomes and stratum, to increase statistical power and precision of the impact estimate (following EEF guidance). Effect sizes will be standardised, and confidence intervals will be reported. We will conduct separate sub-group analysis to understand heterogenous effects by setting type (PVI vs Maintained), EYPP eligibility, quartiles of baseline outcome measure, EAL status of children and training status of practitioners by DfE (Yes/No), the data for which will be collected as part of the practitioner survey.

If we observe statistical imbalance at baseline between the treatment and control groups with respect to characteristics likely to be related to outcomes (e.g. other interventions that may impact the outcomes), we will carry out a multilevel analysis of covariance (ML-ANCOVA) as a secondary model to verify if the findings are robust. This model will explore the impact of controlling for these additional characteristics prior to treatment, alongside a simplified model based on differences in outcome between treatment and control groups. This will indicate how varying the model specification affects the likelihood of detecting impact. A full statistical analysis plan would be developed during the project, and we are committed to working with the approaches set out within the EEF's analysis guidance.

Secondary analysis

For the secondary outcome analysis we will use post-test scores on the FOCUS-34 to directly compare between intervention and control groups. Again, we will use a multi-level model, to take into account the nested structure of the data. Further detail will be provided in the SAP.

For both primary and secondary outcomes, the impact will be presented as a Hedge's g effect size, accounting for the clustering of pupils in settings with 95% confidence intervals (Hedges, 2007).

Sub-group analyses

Sub-group analysis will be conducted for EYPP and for setting type (PVI v Maintained), exploring if Early Talk Boost has a differential impact depending on a children's Pupil Premium (EYPP) status or depending on the type of setting. To analyse these impacts by EYPP status and by Setting Type status, a similar model to that described above will be estimated. To explore the effects of EYPP we will add a binary indicator of EYPP status interacted with the treatment allocation indicator and run the model using a sample that only includes children with EYPP status. To explore the effects of setting type, we will add a binary indicator of Setting Type interacted with the treatment allocation indicator.

Further detail will be provided in the SAP. The statistical analysis will follow the most recent EEF guidance and will be described in detail in a statistical analysis plan, which will be

prepared within three months of randomisation. All analysis will be ITT, where children are analysed as randomised.

Analysis in the presence of non-compliance

The primary analysis would capture an intention to treat effect to minimise bias by preserving the balance afforded by randomisation. It will be important to work with the delivery team to ensure the engagement of settings to minimise the extent of non-compliance. Non-compliance will be examined through the process evaluation. Additionally, practitioners will report how many sessions they have completed, as well as how many children attended each session.

We have also discussed with the delivery team the indicators to include in a binary compliance variable. This compliance variable will be used in a CACE analysis. Children will be considered compliant if they attend at least 21 sessions (7 weeks out of 9 weeks of intervention or 78% of the sessions). Settings will be given a guideline that a session will only be counted if at least 2 children attend the session, and a setting will be considered compliant only if it has delivered 21 such sessions. Therefore, we will conduct CACE analysis only with the child-level binary compliance indicator, as it is assumed that if a setting is non-compliant, the children would also be non-compliant, however, the reverse is not true. Whilst the ITT will provide information on the effect of being assigned to treatment, the CACE analysis will complement this with information on the effect of receiving treatment.

Missing data analysis

We anticipate that a proportion of pupils may attrit before the study is completed or for whom the endline measure of primary or secondary outcome is not available as they are unavailable for testing – which could potentially introduce imbalance with respect to those who do not attrit and complete the endline tests in the intervention and waitlist control group. That may impact the interpretation of the results if the attrition is non-random. Two factors will be looked into while analysing missing data) the extent of missingness and ii) the patterns of missingness. We will firstly check if the missing rate is below 5%, as a missing rate of 5% or less, would not typically bias the primary impact estimates, regardless of the pattern of missingness (Schafer, 1999). If the missing rate is more than 5%, we would study the patter of missingness.

To find out if there is differential attrition based on treatment assignment, we will conduct balance checks between the children who attrit and the pupils who remain in the study, on the following covariates:

- Quartile of their screening results
- Baseline measure of primary outcome
- EY PP eligibility
- Age, gender and ethnicity

If there is a statistical imbalance between the two groups, we will note this in our interpretation of the findings. We will also adjust our power analysis accordingly. We will perform **sensitivity analysis** using listwise deletion in the case of missing data.

Implementation and process evaluation (IPE) design

Research questions

The aim of the IPE is to establish fidelity in implementation, identify the factors influencing implementation that may facilitate/hinder impact at the child-level (including wider national/policy contexts), and to inform and contextualise the quantitative findings. We will examine the delivery of and engagement with the Early Talk Boost programme. In line with the Theory of Change, we will examine the experiences and behaviour of the three key stakeholders: practitioners, children, and parents, as well as interviewing the delivery team.

Using a mixed-method design, the IPE will explore the following dimensions specified in the EEF IPE guidance (2022): fidelity, adaptation, dosage, quality, reach, responsiveness, perceived impact, context/moderators, programme differentiation, monitoring of control group, cost and any possible mediators. We present the IPE research questions under the broad headings of fidelity, implementation and outcomes here but a full mapping of IPE dimensions to research questions and methodology is provided in Table 3 at the end of the IPE section.

Fidelity

To establish fidelity in implementation, we will look at whether the programme is delivered as intended and examine what fidelity and adherence to the intervention means for Speech and Language UK and how well settings have achieved this. This will be explored using the following research questions:

- IPE RQ 1: Can settings deliver with fidelity (i.e., in line with the programme manual) and within routine practice? Elements to explore include:
    - Content (i.e., what was covered in the sessions?)
    - Frequency (i.e., how often were the sessions delivered?)
    - Duration (i.e., did settings deliver sessions for longer/shorter than specified? Were sessions delivered over the intended number of weeks)
    - Coverage (i.e., which sessions were delivered? Are sessions being delivered to the appropriate children, i.e. those who are behind in their communication and language? Within different types of setting?)
    - Adaptations (i.e., do settings adapt any aspects of the programme? What are the reasons for the changes?)
- IPE RQ 2: What does business as usual look like for intervention and control settings, including participation in other language and communication programmes?
- IPE RQ 3: How do practitioners engage with the training and the follow-up sessions? Which practitioners are selected to attend training and deliver the programme?
- IPE RQ 4: Is the intervention feasible for practitioners to engage in and implement as intended? Does this vary by setting type?
- IPE RQ 5: Do children engage with the sessions and the books?

- IPE RQ 6: Are there any considerations for future delivery of the intervention in the current intervention settings or if the intervention were rolled out on a larger scale? What adaptations would be required?

Implementation

We will consider the wider range of issues which affect implementation including the necessary conditions for success and barriers to successful implementation using the following questions:

- IPE RQ 7: What are the potential barriers and facilitators for delivery of the programme? Do these vary by setting type?
- IPE RQ 8: Are there any specific facilitators/barriers for different groups, e.g., children from disadvantaged backgrounds/EAL?
- IPE RQ 9: How are practitioners and settings engaging with parents/carers of children participating in the intervention?
- IPE RQ 10: What, if any, are the costs incurred in delivering the programme, such as staff time or direct costs?
- IPE RQ 11: Are there any unintended consequences or negative effects?

Outcomes

We will explore practitioners' expectations or perceptions around outcomes. These expectations or perceptions are also likely to affect their level of engagement with the programme, and therefore its impact and effectiveness. We will explore the following potential outcomes identified by the Theory of Change model:

- IPE RQ 12: Is there evidence among practitioners of increased understanding, knowledge, and/or confidence to support children's early language development?
- IPE RQ 13: Are practitioners using strategies learned and/or programme activities to support children's language and communication during the intervention? Do they use these strategies/activities outside of the Early Talk Boost sessions?
- IPE RQ 14: What are practitioners' perceptions of any impacts on children's language and communications skills, and/or interactions with others? Does the level of perceived impact differ for different groups, e.g., children from disadvantaged backgrounds/EAL?
- IPE RQ 15: How are practitioners using information from the Tracker to identify children's language needs and inform targeted support for children's language and communication needs, including referral to their specialist services if necessary?
- IPE RQ 16: Are there any changes in the practice of control settings during the trial?

Research methods

Drawing on the EEF IPE guidance (Humphrey et al., 2019; EEF, 2022), we will use a multiphase design to examine the research questions, based around triangulation of mixed methods. This includes:

- Two IDEA workshops and reviewing programme materials; two set up meetings.
- Observations of online training and two online support sessions with practitioners.
- Online or paper surveys for practitioners (baseline and post-treatment).
- Setting visits to eight intervention settings towards the end of the programme to observe delivery and conduct interviews with practitioners, year group/language/Early Years leads and the setting manager.
- Interviews with up to ten parents at settings that are delivering the parent workshop, recruited via settings sampled for setting visits in the first instance.
- Interviews with up to five settings (setting managers) and up to five practitioners that have dropped out.
- Five interviews with the delivery team at the end of delivery.
- Analysis of data collected by the delivery team, such as training attendance and feedback or requests for support/guidance..

The issue of saturation has been considered in the choice of numbers of interviews, case studies and workshops. Studies in the qualitative research literature have found that most themes were identified in the first 5-6 interviews (Morgan et al., 2002) and that very little new information was gained after 9-17 interviews (Hennnik and Kaiser, 2022). Drawing upon these findings, we have set out the factors that have informed our choice of interview numbers for each element as relevant. Interviews will be digitally recorded with the agreement of participants and transcribed verbatim.

In the **IDEA workshops**, our own and Speech and Language UK's team will explore the intervention as part of an initial session shortly after set-up (January 2023) and then another late in the delivery period to capture any further changes (June/July 2024). Building on the set-up meetings, we will: review the TIDieR framework and theory of change; examine training/delivery materials; re-visit evidence about the interventions. This will enable us to identify and agree with the delivery team the key expected outcomes, the mechanisms for this change, and an appropriate measure of compliance. The outcomes will be measured in the impact evaluation using standardised tests, but the IPE will also explore perceived outcomes, as well as investigating evidence for the mechanisms of change and factors that may influence these. Again, data for an agreed measure of compliance will be recorded as part of the impact analysis but the IPE will also provide opportunity to collect wider information around compliance, e.g., what this looked like and what factors affected this.

As part of the training stage, we will **observe one each of the two half-day online training sessions for practitioners and two online support sessions**. We will observe delivery for both Stronger Practice Hubs, e.g., one training session with a Hub 1 cohort and one with a Hub 2 cohort, and one online support session from each Hub cohort. This will give the evaluation team further understanding of how the intervention is intended to be delivered and any early issues or contextual factors identified as possible moderators or mediators. This information will inform the development of setting visit interviews and practitioner survey questions around fidelity/adherence, dosage, quality, reach, adaptation and possibly also context/moderators and mediators. The support sessions take place during delivery, ideally while practitioners are implementing the intervention, and may provide insights into factors that are facilitating or hindering implementation (context/moderators) and how components of the intervention are being delivered (quality, fidelity/adherence). As the aims of the

observations is to inform development of research instruments, rather than provide a representative picture of training delivery and practitioner response, we have chosen to observe one each of training and support sessions for each of the Stronger Practice Hub areas, in order to reduce burden and interference with delivery while gaining a snapshot for each area.

We will create semi-structured observation frameworks that incorporate the AEIOU model (Robinson et al.,1991). This is an ethnographic approach which organises information into five elements: Activities, Environments, Interactions, Objects and Users. This will enable evaluators to collect and organise rich data efficiently, while also being flexible to variations in training or personalised support and wider context. As the aim of the training and support observations is to inform the development of research materials, the frameworks will be used to collect a small sample of qualitative data where emerging themes can be identified. We will also **review the learning materials**, e.g., manual. This will enable the evaluation team to develop an understanding of the intervention, prepare our observation, interview and survey instruments, and inform observations and interviews.

The **survey of practitioners,** (across all settings) will be developed using online survey software, SNAP, which allows completion on mobile devices via hyperlink or QR code and can be provided in a paper format if needed. At the start of the 2023/2024 academic year, before randomisation, all settings will be asked to nominate two practitioners who work with 3-4-year-old children to receive training if their setting is allocated to the intervention group. These practitioners will be surveyed at baseline and endline, and comparisons made between the intervention and control groups. The baseline survey will take place in October-November 2023 as most children will have started for the term and classes/rooms will be established and somewhat settled-in. The endline survey will take place in end of April-May 2024 after delivery of the Talk Boost intervention has finished and practitioners are able to reflect on this process. It is likely that there will be some turnover of staff between the start of the academic year and the training in January 2024 so we will ask the delivery team to provide a list of practitioners who have participated in the training so that we can update the intervention group respondents list for the endline survey, i.e., add/remove respondent. There is also likely to be some turnover between the baseline and endline survey more generally although delivering the survey in April-May rather than at the end of the year should mitigate that slightly. For control group practitioners, where we find that email addresses are no longer valid or a respondent/setting lets researchers know that they are no longer working with 3-4-year- children then we will contact settings to ask for contact details for alternate staff working with that age group. Where respondents have dropped out or been added, it will not be possible to track change over time but it will still be possible to make comparisons at single time points between practice at intervention and control settings.

If there are 132 settings in the trial, then 264 practitioners would be invited to participate in the surveys. Response rates can be relatively low for practitioners in Early Years settings but if only 38% of practitioners respond that would still be 100 responses. However, we would not want to go below 25% of the sample (66 respondents) and would want coverage of staff from at least 40 settings. It would also be important to have respondents from each Stronger Practice Hub region. Ideally, this imbalance would not be larger than 60%/40%. We would monitor this during fieldwork and if one area was underrepresented, we would contact settings in those areas that had not yet responded to ask them to encourage relevant staff to respond, or we could also approach Stronger Practice Hubs to encourage settings in their area to respond.

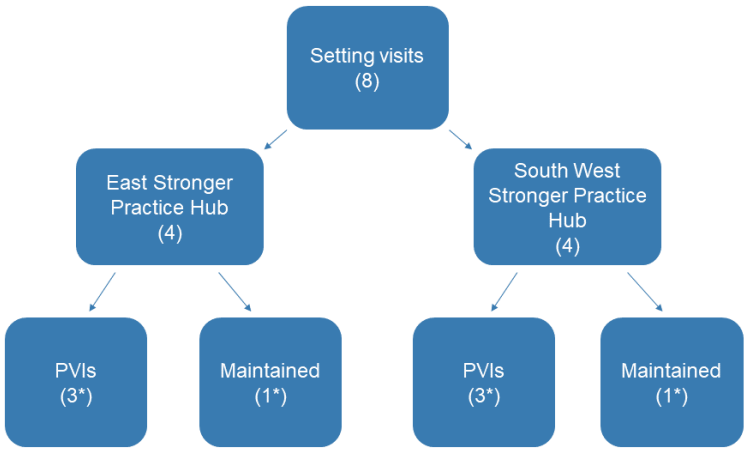
The baseline survey will capture usual practice prior to randomisation, staff characteristics and experience, and information on broader approaches to supporting language and communication. The endline survey will repeat questions about language and communication practice, and include questions on wider context, e.g., changes in practice at setting level outside of the intervention, and costs. For treatment settings, this will cover experiences of taking part, staff time and resources required (to inform the cost-per-child estimate). For control settings, this will include questions around business-as-usual costs.

For the questions around practice to support language and communication, we will draw upon a survey co-developed with researchers at the University of Oxford, as part of the Coaching Early Conversation, Interaction and Language (CECIL) study (Dawson, Huxley & Garner, 2022), designed to capture confidence, knowledge and practice to support children's language and communication development. This survey would be developed in consultation with the delivery team to ensure that items identifying practice developed/learned through participation in Early Talk Boost, as well as wider practice that may be influenced indirectly. These questions will aim to capture confidence, knowledge and practice supporting children's language and communication development. For both confidence and knowledge or practice questions, respondents will be asked to rate their current level using a five point scale.

In addition to this, intervention group practitioners will be asked questions around fidelity of Early Talk Boost sessions, and control group practitioners will be asked questions around any support they provide for children who are struggling slightly with communication and language, i.e., at a slightly earlier stage than their peers. This should allow the evaluation to identify any change in practice or compensatory behaviour in control settings.

Eight setting visits will allow detailed qualitative exploration of delivery. The sample frame will include an even split of settings across the two Stronger Practice Hub areas (East , and South West), and will include a mix of PVI and maintained settings (see figure 2). The balance of PVI to maintained settings will be informed by the numbers of each recruited to the study by the delivery team and by feedback from the Stronger Practice Hubs regarding the typical prevalence of these setting types in their area. Within this frame, selection of settings will also be informed by geographical context (urban/rural), number of 3-4-year-olds at the setting, proportion receiving the Early Years Pupil Premium (EYPP), whether the setting is part of a chain and potentially also compliance. While it is not possible to account for all these factors with the current sample size, where possible we would seek to ensure that, for example, not all settings visited were located in cities. The delivery team are collecting information during recruitment regarding current number of 3-4-year-old children, 3-4-year-old children receiving the Early Years Pupil Premium (EYPP) and setting type. Data regarding number of 3-4-year-old children and EYPP will be updated during data collection by the evaluation team at the start of the academic year. If the baseline survey indicates key differences in pre-intervention language and communication practice, this could also inform selection of settings within the sample frame. We believe that eight setting visits should be sufficient to explore some key differences between settings as well as some variation in other characteristics without causing burden on too many settings, and would allow for interviews with 8-16 participating practitioners, 8 setting managers/Early Years Leads/Room managers and 8 observations. Setting staff will receive an incentive of £15 for participating in an interview. If elements of a setting visit are unable to take place for one or two of the settings then there should still be sufficient data for each element, e.g., 6 manager interviews, for analysis to capture the main themes.

Figure 2: Sampling framework for setting visits



*Estimated values. Sample numbers will be decided after reviewing the distribution of setting type among settings participating in the trial and guidance from Stronger Practice Hubs on distribution in their area.

Other considerations

- Urban/rural
- Size: number of children aged 3-4 years
- Proportion of children receiving Early Years Pupil Premium
- Whether setting is part of a chain
- Whether delivering the Early Talk Boost parent/carer workshop
- Compliance
- Pre-intervention language and communication practice

In each setting visit, we will observe Practitioners delivering a Talk Boost session, and use these data to inform assessment of fidelity and shape questions for follow-up interviews. The observations will be focused on feasibility of delivery of the intervention and child engagement and will not be a rating of compliance. We will create a semi-structured observation framework using the AEIOU model, in consultation with the delivery team, that will allow us to identify key elements of the intervention, as well as observing how children are engaging with the delivery. This qualitative data collected during observations will be incorporated into the analysis framework with interview data collected during the setting visit. Where possible, we will interview the practitioner who is delivering the intervention, the other practitioner that has been trained and a senior leader/setting manager or a language/Early Years lead. We anticipate an individual interview with senior leaders/setting managers and will interview practitioners separately to ensure open/honest discussions. Interviews may explore — training received; materials; workload/time requirements of Talk Boost; costs incurred; facilitators/barriers to implementation; adaptations and reasons for these (including views on how children with EAL respond); how Talk Boost compares to their usual support approach for children struggling with language and communication; children's outcomes; and suggested improvements. Interviews will last around 30-45 minutes and will be at flexible times as staff may have limited availability due to the need to maintain staff:child ratios. Where it is not possible to talk to a member of staff during the visit, we would offer to conduct the interview at a different time via phone or video-conferencing and at a time convenient to the staff member. When arranging setting visits, we would let the settings know the elements we are requesting, e.g., observations, interviews, and the setting could then suggest timings for these as suits them, e.g., attending morning and afternoon sessions separately across two days. Evaluators would also seek to be responsive and adaptive on the day so as not to interfere with the setting's delivery.

Interviews with up to ten parents or carers. As the parent workshops have not been classified as a core requirement of the intervention, this aspect is a scoping exercise to capture some additional information around how settings are currently engaging with parents. Interviews may explore parents/carers' experience of and views on the workshop, Tizzy's Talking Tips and the Jake and Tizzy books, enablers and/or barriers to using these, how they

are supporting their child's language development more generally, and their confidence around this. We would aim to speak to ten parents or carers at settings that are delivering the workshop with parents/carers. Parents would receive an incentive of £5 for participating in the interview. We would seek to recruit these as part of the setting visits. Ideally, we would include this among the criteria in the sampling frame for setting visits. However, it is not possible to know which settings have delivered the parent workshop until they start delivering the programme, and settings may not all start delivering the programme at the same time. The delivery team may pick up this information during the support visits but that also depends on when they speak to settings, as support visits may take place before the parent workshop would be delivered. For this reason, we will start recruiting to parent interviews through setting visits initially and then expand to other approaches if needed.

During recruitment of a setting visit, we would ask whether or not that setting was delivering the parent/carer workshop, and whether the setting would be willing to ask any parents or carers of children participating in the evaluation whether they would be willing to participate in a short interview with the evaluation team. We would be led by the setting as to whether they felt comfortable approaching parents/carers and how they would prefer to facilitate these interviews, e.g., passing on contact details or arranging for interviews to take place at the setting. We would speak to one or two parent/carers at most per setting to allow for interviews across a few settings in order not to overburden one setting. Parent/carer interviews may need to take place separately by phone/video-conferencing if it is not possible for these to take place during the visit. If it is not possible to obtain the target number of parent interviews as part of setting visits, the evaluation team will ask the delivery team if they are able to suggest settings that are delivering the parent workshop or not by drawing on information collected during the support sessions. If this approach is not possible, then the evaluation would consider emailing settings directly to ask whether they are currently delivering the parent workshop and whether they would be willing to facilitate interviews with parents/carers of participating children. We will aim to explore parents/carers' experiences of engaging with and delivering Talk Boost in the home and, where relevant, their experience of participating in the parent workshop.

Follow-up interviews with up to 5 practitioners and up to 5 leads at settings that have dropped out of the study. Where we have been informed that settings or practitioners have dropped out and where appropriate, we would seek to invite them to take part in short 10–-15 minute interviews. We would take guidance from the delivery team and, for practitioner interviews, speak to setting leads regarding initial reasons for withdrawal and whether there were any reasons why it would be inappropriate to make contact. Interviews will be shorter to reduce burden on participants and motivate them to take part. At the setting level, we will explore barriers to continuing with the intervention and any potential solutions that could be offered. At the practitioner level, we would explore reasons why individuals were no longer involved in delivering the programme and any potential barriers to participation at an individual level. Five interviews should be sufficient to enable identification of key themes. We will alternatively use a short email form if interview uptake is low.

We will also conduct **up to 5 telephone interviews with the Speech and Language UK delivery team** towards the end of intervention delivery. This would include interviews with those involved in managing and designing the intervention, as well as with a staff member delivering support sessions to settings for each of the two Stronger Practice Hub areas. It should be sufficient to interview one staff member delivering support sessions for each of the areas, as staff managing the delivery team will have an overview of the key issues which have

been fed back to them by the wider team. The interviews will explore delivering training, setting engagement and participation, and enablers/barriers to successful implementation and cascading of the programme.

Alongside this, we will also **collect data on fidelity** as defined in the impact analysis, including practitioners' attendance at training and child attendance at Early talk Boost sessions. We will also review **programme data** shared by the delivery team. This will include feedback on the initial training and queries/support requests from settings. This allows identification of issues raised by settings and how these were resolved or support by the delivery team. Please see the discussion of 'Delivery information' in the analysis section below.

Analysis

Qualitative and quantitative data from the case studies, interviews, surveys and delivery information will be integrated in the final analysis to create a rich picture of delivery, participant experiences and setting staff views to investigate mechanisms of change set out in the Theory of Change model and inform impact analysis. The different research methods will capture different perspectives and different levels. Survey and delivery data will provide high level overviews and broader patterns across participating settings, whereas interviews and observations provide in-depth information regarding a limited number of settings/individuals that can offer insights into what factors may contribute to broader trends. Information from different sources will be compared to identify where findings are supported across multiple sources, and where differing findings indicate that closer analysis is needed to understand minority perspectives/experiences, differences in perspective/experience by stakeholder, or differences between reported and observed/measured experience. These findings and data will be synthesised to create robust and nuanced conclusions that test mechanisms of change set out in the Theory of Change model and inform impact analysis

Framework analysis

A key tool for synthesis of findings across multiple sources is the 'Framework' approach. We will analyse qualitative data using 'Framework', drawing themes and messages from an analysis of interview transcripts, observations of delivery, and other materials collected by evaluation and project teams, such as requests for support. Information regarding some setting or individual characteristics will also be incorporated into the framework to provide context, such as job role or Stronger Practice Hub area. Codes will be used to enable the identification (and filtering) of qualitative data from the same setting and similarly a common descriptor will be used for each type of qualitative data (practitioner interview, parent interview, observation etc).

Framework is an excel-based qualitative analysis tool that ensures that the analytical process and interpretations from it are grounded in the data and tailored to the research questions. It is designed to ensure a systematic and consistent treatment of all units of data (e.g. transcripts of interviews) and allows for the analytical framework to be refined and modified in the early stages of its use. We would use a mix of deductive and inductive approaches to analysis and coding. Before analysis begins, the framework will be set up with themes drawing upon the implementation dimensions and the IPE research questions, i.e. using a deductive approach. However, during analysis or data collection, evaluators may identify additional issues or questions that could contribute useful insights so these would be incorporated into the framework, i.e. an inductive approach.

Information from each interview or observation will be entered as a separate row in the Framework and represents one perspective. Framework allows full within case analysis (looking in detail at each individual row) and between case analysis (comparing individual rows and groups of rows), as well as thematic analysis. This will facilitate analysis at a setting level to understand broader impacts and experiences within an individual setting context, as well as being able to compare individual experience across a specific characteristic, e.g. role, setting type. We would analyse individual experience by a specific characteristic first to identify the main themes and then analyse by setting to understand how context affects experiences related to these themes. Verbatim quotes and evaluator notes are included in the Framework under relevant themes or questions. The context of the information is retained, and the transcript location is noted, so that it is possible to return to a transcript to explore a point in more detail or to extract further text where a verbatim quotation is particularly long. Organising the data in this way allows us to compare the full range of experiences and accounts and patterns across different groups of people.

Statistical analysis

Quantitative analysis of the practitioner survey at both baseline and endline will include basic statistical tests to identify changes in practice or behaviour, using descriptive frequencies and cross-tabs, as well as inferential statistics.

Data will be analysed to compare intervention and control groups and controlling for change over time (if we have sufficient repeated measures data). This will either be achieved through a creation of a change score from pre-test to post-test and then an application of a t-test or non-parametric equivalent such as a Mann-Whitney to compare the groups, or by comparing the post-test scores between groups, controlling for pre-test responses. This would be achieved using an ANCOVA model or equivalent non-parametric design. This approach would also allow addition of other covariates such as Stronger Practice Hub. If due to churn of staff the proportion of respondents completing both pre- and post-test is low then we will also compare the post-test scores between groups without the addition of the pre-test covariate.

Further details will be provided in the SAP.

These findings would then be triangulated with findings from qualitative and monitoring data to explore what factors may contribute to the findings or highlight areas for closer consideration where findings from different sources presenting conflicting pictures.

Delivery information

Programme information provided by the delivery team would be used to provide broader insights into delivery that can be used to enrich other analyses. As part of the compliance measure for the impact analysis, the delivery team will share data on practitioner attendance at training and support sessions, practitioner delivery of Early Talk Boost sessions, and children's attendance at Early Talk Boost sessions. This will enable the evaluation team to identify settings where compliance has been low so this characteristic can be incorporated into analysis of staff interview and/or practitioner survey data. Settings could be grouped by compliance level and this characteristic can be incorporated into framework analysis or as an independent variable in statistical analyses. Where low compliance is identified at an early stage, the evaluation team may consider including at least one low compliance setting in the setting visits sample, although there would need to be consideration of additional burden as these settings may be struggling more generally, e.g. staff shortage. It is also possible that

some low compliance settings may then go on to withdraw from the programme or study. Analysis pertaining to individual children regarding compliance in settings will be undertaken as part of the impact analysis.

Information around support requests to the delivery team, such as email queries or topics raised in support sessions, can be thematically analysed to identify key areas of support and evaluators can explore to what extent this reflects themes emerging from setting visits, as well as giving an indication of support needs across the wider sample. Participant feedback from training sessions will be shared to provide initial insights into practitioner views on the training which can be explored and/or compared with findings from staff interviews. Finally, the delivery team will be asked to provide information about the current price of the intervention for settings outside of the trial, whether they have made any changes to the intervention during delivery that would mean that the price would be different for the intervention as it has been delivered during the trial, whether they anticipate any change to the price for future delivery, and whether they anticipate any change in the price of the intervention if it were to be delivered on a larger scale. Further detail is provided in the cost evaluation section of the protocol.

Table 3 presents an overview of how the IPE research methods and questions feed into the IPE dimensions identified for this study, including a brief description of sampling, participants/data sources, data collection and analysis information.

Table 3: IPE methods overview

IPE dimension	RQ addressed	Research methods	Data collection methods	Sample size and sampling criteria	Data analysis methods
Fidelity	RQ1, RQ3, RQ4, RQ5, RQ8, RQ9, RQ10.	Theory of Change development	Two IDEA workshops	N.A.	Theory of Change
		Material review	Reviewing intervention materials	N.A.	Literature review, thematic analysis
		Observations of training	Semi-structured observations	2 training sessions & 2 support sessions: 1 each with each Stronger Practice Hub cohort	Observation framework, thematic analysis
		Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs
		School case studies	Semi-structured observations, semi-	8 settings (sizes/types, SPH, high/low compliance), 8 setting	Teaching observation framework, Framework

			structured interviews	managers/EY Leads, 16 practitioners, 8 intervention sessions	analysis, thematic analysis
		Delivery team interviews	Semi-structured interviews	Up to 5 members of S&L UK delivery team (programme managers & developers, delivery staff from each hub)	Framework analysis, thematic analysis
		Monitoring Information	Analysis of intervention data collected by S&L UK	All programme	Thematic analyses, frequency counts, average cost per child
Adaptation	RQ1, RQ3, RQ4, RQ5, RQ6, RQ7, RQ8, RQ9, RQ10, RQ12.	Theory of Change development	Two IDEA workshops	N.A.	Theory of Change
		Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs, t-tests/Mann-Whitney, ANCOVA/non-parametric equivalent
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis
Dosage	RQ1, RQ3, RQ4, RQ5, RQ8, RQ9, RQ10	Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis

		Monitoring Information	Analysis of intervention data collected by S&L UK	All programme	Descriptives: frequencies, cross-tabs,
Quality	RQ1, RQ3, RQ4, RQ5, RQ8, RQ9, RQ10, RQ12.	Observations of training	Semi-structured observations	2 training sessions & 2 support sessions: 1 each with each Stronger Practice Hub cohort	Observation framework, thematic analysis
		Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis
		Delivery team interviews	Semi-structured interviews	Up to 5 members of S&L UK delivery team (programme managers & developers, delivery staff from each hub)	Framework analysis, thematic analysis
Reach	RQ1, RQ3, RQ4, RQ5, RQ8, RQ9, RQ10.	Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs,
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis

		Monitoring Information	Analysis of intervention data collected by S&L UK	All programme	Thematic analyses, Descriptives: frequencies, cross-tabs, average cost per child
Responsiveness	RQ1, RQ3, RQ4, RQ5, RQ8, RQ9, RQ10.	Observations of training	Semi-structured observations	2 training sessions & 2 support sessions: 1 each with each Stronger Practice Hub cohort	Observation framework, thematic analysis
		Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs,
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis
		Delivery team interviews	Semi-structured interviews	Up to 5 members of S&L UK delivery team (programme managers & developers, delivery staff from each hub)	Framework analysis, thematic analysis
		Monitoring Information	Analysis of intervention data collected by S&L UK	All programme	Descriptives: frequencies, cross-tabs,
Programme differentiation	RQ2, RQ4, RQ6, RQ7, RQ8, RQ9, RQ11, RQ12.	Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis

Monitoring of control group	RQ2, RQ13, RQ14, RQ15, RQ17.	Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs,
Perceived impact	RQ13, RQ14, RQ15, RQ16, RQ17.	Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs, t-tests, ANOVA, regression
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis
		Delivery team interviews	Semi-structured interviews	Up to 5 members of S&L UK delivery team (programme managers & developers, delivery staff from each hub)	Framework analysis, thematic analysis
Cost	RQ1, RQ2, RQ4, RQ6, RQ7, RQ8, RQ9, RQ11, RQ12, RQ17	Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs,
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis
		Monitoring Information	Analysis of intervention data collected by S&L UK	All programme	Thematic analyses,
		Withdrawal interviews	Interviews with practitioners / settings that withdrawn from the study	10 practitioners	Framework analysis, thematic analysis
Context/moderators	RQ1, RQ2, RQ3, RQ4,	Observations of training	Semi-structured observations	2 training sessions & 2 support sessions: 1 each with each Stronger Practice Hub cohort	Observation framework, thematic analysis

	RQ5, RQ8, RQ9, RQ10, RQ11, RQ12, RQ13, RQ14, RQ15, RQ16, RQ17	Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs, ANCOVA/non-parametric equivalent
		School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis
		Parent interviews	Semi-structured interviews	10 parents/carers from setting visits (lower/higher engaged)	Framework analysis, thematic analysis
		Withdrawal interviews	Interviews with practitioners / settings that withdrawn from the study	Up to 5 practitioners, Up to 5 withdrawn settings	Framework analysis, thematic analysis
		Delivery team interviews	Semi-structured interviews	Up to 5 members of S&L UK delivery team (programme managers & developers, delivery staff from each hub)	Framework analysis, thematic analysis
Mediators	RQ2, RQ8, RQ9, RQ11, RQ12, RQ17	School case studies	Semi-structured observations, semi-structured interviews	8 settings (sizes/types, SPH, high/low compliance), 8 setting managers/EY Leads, 16 practitioners, 8 intervention sessions	Teaching observation framework, Framework analysis, thematic analysis
		Practitioner survey	Online questionnaires (baseline and post-treatment)	264 practitioners: 2 at each setting (all)	Descriptive frequencies, cross-tabs, ANCOVA/non-parametric equivalent
		Delivery team interviews	Semi-structured interviews	Up to 5 members of S&L UK delivery team (programme managers & developers, delivery staff from each hub)	Framework analysis, Thematic analysis
		Parent interviews	Semi-structured interviews	10 parents/carers from setting visits (lower/higher engaged)	Framework analysis, thematic analysis

Cost evaluation design

Following the EEF principles of cost evaluation in the latest guidance we plan to follow the prescribed ingredients method, gathering data from the programme providers and setting staff, to arrive at a cost per pupil per year for engaging in the intervention. To do this we will collect information on the pre-requisite, set-up, and ongoing costs to settings, as well as information on staff time spent related to the delivery of the Early Talk Boost programme.

We will use multiple sources and methods to determine these to provide a full evaluation of costs involved. We will do this via a bespoke survey, designed in house at IES, and completed by settings managers, and via in depth qualitative interviews with the setting staff, and the delivery team.

In order to design the most appropriate and effective survey we propose to complete an ex-ante costing exercise with the delivery team where we will establish the main usual expected costs. We will then design our practitioner survey using these answers to efficiently capture the expected costs in a pre-populated way. This will allow us to include drop down options in the survey which participants can select answers from which reduces burden both on the participant and also on the later analysis. We will also allow optional free text responses from those who wish to add more detail, and for more bespoke responses. We will analyse this data alongside the qualitative data from the in-depth interviews. The in-depth interviews will be exploring staff time used to run the intervention and general reflections on costs and perceived value for money and affordability.

Following the EEF guidance we will be calculating and presenting the total cost per setting for the intervention as implemented over three consecutive years, and the cost per-child-per-year. The costs will be estimated for the programme as it was implemented during the trial and estimated using market values (i.e., not including any subsidies provided by the EEF for the purposes of the trial). Published unit costs will be utilised where possible, for example salary costs. Costs will be valued as per the year of analysis (expected 2024). Depending on response rates and data collected we may do some sensitivity analyses to account for any uncertainty in the costing estimates and to estimate the cost impact of variations to implementation delivery.

Ethics and registration

The trial will be designed, conducted and reported to CONSORT standards (http://www.consort-statement.org/consort.statement/) and registered on the ISRCTN registry (http://www.controlled-trials.com/).

The evaluation has been approved by IES ethics committee after careful consideration and a meeting between the committee and IES team members. All of the recruitment documents were jointly developed and agreed by Speech and Language UK and IES, with comments from EEF.

The delivery team will share setting level information sheets with the settings detailing the study. Ethical agreement for participation within the trials will be provided by the setting manager by providing informed consent to the delivery team via signing an MOU that covers

information about the study, the respective responsibilities of setting, evaluation team and programme delivery team and the ways in which the data will be handled under GDPR regulations. The MOU also covers future data archiving in the EEF archive. Before requesting settings to share the pupil data, IES will ask the settings to distribute a parent information sheet and parent withdrawal forms. This will enable parents to withdraw their child from the study if they wish to do so. Settings will not share child data for the children whose parents withdrew them at this stage. Settings and parents have a right to withdraw from the trial at any point with no detrimental consequences. Setting staff will give informed consent to be observed and to take part in interviews. Setting staff will also be given the opportunity, if they request it, to view a transcript of their interview. Details about viewing their transcript and the right to withdraw their data from the evaluation following participation in the study will be provided in an information sheet for setting staff. Setting staff will have 30 days after their interview to request a copy of their transcript and 30 days after receiving their transcript for the data to be withdrawn. In addition to the information sheet, verbal information will be provided prior to any research encounter, and ethical consent will be confirmed. Verbal information will make clear the purpose of the research, how the data will be used, when the data will be deleted and how to withdraw data. Associated privacy notices will also be issued where appropriate.

Data protection

We recognise the importance of data protection and are committed to complying with the Data Protection Act 2018 and the GDPR. Our approach involves collection of child level data, individual level survey data, setting level data and in-depth interviews. The evaluation team at IES and the delivery team at Speech and Language UK will collect personal data about practitioners, settings, children and parents at different times throughout the evaluation. The evaluation team's legal basis for processing personal data is legitimate interest and this is necessary scientific research purposes. We have developed a data flow document that includes how and why information will be shared between evaluation teams, the delivery team and EEF. Speech and Language UK are the independent controller for the data they are collecting as part of recruitment of settings and for programme delivery. Their legal basis for processing this data is legitimate interest. IES has produced a DPIA for this project, which indicates oversight by the IES DPO with ongoing review.

For the purposes of conducting the evaluation to assess the impact of Early Talk Boost, IES will become data controller of personal data of setting staff and children obtained from settings. They may share personal data with the delivery team as well as trusted processors such as academics, test administrators, transcribers and research assistants solely for the purposes of proper delivery, management and evaluation of the project. Speech and Language UK and IES will sign a data sharing agreement that outlines all of the data that will need to be shared between the two organisations.

We will ensure that participating settings and parents understand the evaluation requirements. A privacy notice will explain how information collected would be used and stored, communicate to participants and their right to withdraw from data processing. All participants interviewed or surveyed for the research will be asked for their consent to indicate that they understand the aims of the research and agree to the interview being recorded and transcribed and given assurance of anonymity.

Settings will sign an MOU at the start of the project clearly laying out the requirements of participation and how data will be used, shared, and stored. Some of the key ethical

considerations that we will take into account while developing the protocol for this project would be minimising burden on settings and children; ensuring settings and parents are able to make an informed decision about taking part.

All participating settings will be required to distribute an information letter and privacy notice to parents of 3- to 4-year-old children in September 2023. This letter and privacy notice will be provided by IES and shared with settings.

The letter and privacy notice will inform parents of the nature of the project, the personal data that will be collected about them and their child and how this data will be processed. It will also give them the opportunity to withdraw from the project if they wish. Providing parents/carers with the opportunity to withdraw their data and the data of their child from the project if they so wish will provide due consideration for their privacy and rights in relation to their data.

Settings should allow parents an initial two weeks to respond if they wish to withdraw from the study. In cases where parents inform the setting that they do not wish to participate in the project, settings should remove data of these children and parents from the data they share with IES.

Beyond the initial two-week period, parents/carers will be able to withdraw their child's participation from the entire study by contacting IES directly at *ETBeval@employment-studies.co.uk.* Parents/carers will be provided with full details on their rights under data protection laws and contact details for the project team in the information letter.

At the end of the project, a pseudonymised dataset will be submitted to the EEF's data archive which is managed by the Fischer Family Trust (FFT). At this point, EEF will become a data controller for the datasets archives after the trail (following the successful completion of internal quality checks by the archive manager) and IES will be a data processor. More details about EEF's data archive can be found here: https://educationendowmentfoundation.org.uk/privacy-notices/privacy-notice-for-the-eef-data-archive. Personal data, interview notes and survey responses will be securely deleted from the IES systems six months after the project is complete (currently estimated to be July 2025). No personal data will be shared outside of the EEA.

Personnel

IES

Seemanti Ghosh, PhD, Principal Research Fellow (Head of Impact Evaluation) is an economist by training and specialises in quantitative evaluation. She will direct the evaluation. She has a methodological expertise in RCTs and her experience of conducting trials spanning over eight years, with a particular focus on cognitive and noncognitive development of children. Her doctoral work investigates the impact of adverse experiences on socioemotional development of children through an RCT. As a postdoc she conducted a large-scale RCT in India where she evaluated a parental engagement intervention for 200 public schools, including 6,000 parents and 1,200 teachers. She joined IES in September from NESTA where she was the Principal Researcher (Evidence and Experimentation) and led design of trials evaluating behavioural interventions. At IES, Seemanti heads our impact evaluation work.

Anneka Dawson, Dphil, Principal Research Fellow, (**Head of pre-16 education and IES Trials Unit)** will provide senior trials advice and expertise. Anneka has substantial expertise in education and family research spanning over 15 years. Anneka has directed three IES projects with EEF as well as leading consultancy work for EEF of an overview of their Early Years work in 2018-2019. She has directed other early years projects recently including CECIL, OVO Foundation and CLPE. She was formerly the senior evaluation manager at the EEF and had responsibility for the Early Years research.

Susie Bamford, PhD, Senior Research Fellow will lead the impact evaluation. Susie specialises in quantitative and statistical research. Susie has substantial trials' experience spanning over 20 years. She completed several years of postdoctoral research in developmental psychology laboratories running RCTs with behavioural and educational outcomes in schools, nurseries, and clinics. Following this she gained three years of educational research experience at NFER where she led an RCT of Code Club and was involved in both the process and impact side of the FAST trial. Over the last four years she has run her own business providing consultancy on the design and delivery of research trials to various clients. Susie has extensive experience designing research for and working with young children and families including working as a Research Fellow on the Comparison of Preschool Parenting Programmes RCT at the University of Southampton.

Clare Huxley, PhD, Senior Research Fellow, will lead the implementation and process evaluation. Clare has research interests in education and is experienced in a range of qualitative and quantitative approaches. Her PhD is in psycholinguistics (University of Edinburgh) and she has knowledge of childhood language processing and acquisition development. Clare managed the EEF Early Years Toolbox pilot study and is currently managing the Flexible Phonics efficacy trial. She also managed the implementation and process evaluation for the CECIL study funded by The Sutton Trust, Esmée Fairbairn Foundation and Lindsell Foundation.

Alexandra (Allie) Nancarrow, PhD, Research Fellow, will manage the evaluation. Allie focuses on research into early childhood education. Her main interests include using a range of quantitative skills to investigate predictors of children's educational achievement and cognitive development. She has experience with large-scale, longitudinal research projects, analysing large secondary datasets, and disseminating research to families, school personnel, stakeholders, and charities.

Jade Talbot, MSc, Research Officer, will lead the process evaluation fieldwork. Jade has research interests in education and life chances and is currently working on the pilot evaluation of Catch-Up Literacy for What Works Children's Social Care as well as supporting two projects for the Youth Futures Foundation to generate understanding what works to support disadvantaged young people into further education and employment. Jade has extensive experience of qualitative research methods, such as learner survey design, semi-structured and in-depth interviews, in-person observations, as well as producing literature reviews and case studies.

Meghna Sharma, MSc, Research Officer, will provide support throughout the project. Meghna recently joined IES after completing a Masters in Social Policy and Research from University College London. She has a range of research experience, including conducting focus groups, semi-structured interviews, and analysing data, and is interested in areas such as education and health. She has also worked as a teacher for children from low-income backgrounds.

Speech and Language UK

Louisa Reeves, Director of Policy and Evidence. Louisa is the project sponsor for Speech and Language UK. Louisa is a speech and language therapist with 30 years' experience of working to support children and young people with speech and language challenges. She has been involved in a number of large-scale projects at Speech and Language UK, developing evidenced solutions for speech, and language challenges. Her work includes leading on the development and evaluation of our Early Talk Boost and Talk Boost KS2 language interventions and the Tots Talking parent intervention for 2 year olds.

Liz Wood, Lead Speech and Language Advisor, is leading on the recruitment and delivery of Early Talk Boost into settings. She is a Speech and Language Therapist and joined Speech and Language UK in 2007. Liz has worked on the set up and delivery of several projects and been involved in the development of Speech and Language UK interventions and resources, as well as training and consultancy crossing all phases of education. Liz is experienced in workforce development, creating and delivering training and workshops. She has worked in partnership with LAs, settings and organisations to develop communication supportive environments and improve practice to support children and young people with speech and language challenges.

Jane Flynn, Monitoring and Evaluation Manager, is providing support throughout the project. Jane joined Speech and Language UK after completing a Masters in Speech and Language Therapy. She has worked on a number of Speech and Language UK's evaluations over previous 4 years covering early years through to secondary.

Risks

Risk	Action to reduce risk	Likelihood	Impact
Timeline for pre-testing	The timeline for the pre-tests will be clearly communicated with assessors and setting staff, including reminders about testing dates/times.	Low	High

Settings not available for training on little notice	Settings will be notified as soon as possible about training dates.	Medium	High
Insufficient settings recruited/retained	The delivery team will work closely with the Stronger Practice Hubs to recruit settings.	Medium	Medium
Insufficient numbers of children recruited/retained	The delivery team will work closely with the Stronger Practice Hubs to recruit settings and highlight the importance of the Early Talk Boost intervention to setting staff. If settings have fewer than 3 eligible children, then those settings will be dropped from the study.	Low	Medium
Settings do not provide data	Responsibilities will be clearly laid out in an MoU. We will work to reduce burdens on settings with clear guidelines on requirements, with long lead in times to help planning. Setting staff will receive a £15 incentive for participating in an interview. A broad team means we can share following up with settings to ensure timely provision of data.	Medium	Medium
Attrition, especially non-completion of post-tests	Requirements of the study, including assessments, will be made clear to settings in recruitment. Ensuring flexibility in the scheduling of assessment and minimising burden for settings where possible. We can keep the control settings engaged by building in touchpoints at least twice before they start the intervention over the next term.	Medium	High
Reluctance of practitioners or parents to participate in surveys and interviews	Parents and practitioners will be made aware of the expectations of the study in recruitment. We recommend £15 incentives for practitioner interviews and £5 parent interviews. Interviews will be scheduled to be convenient for practitioners and parents. Surveys for both parents and practitioners will be short and using SNAP software which is very	Low	High

	accessible. Mobile-friendly format, so that they can complete the survey in preferred platform.		

Table 4: Timeline

Activity	Time	Responsibility
Soft recruitment (project page set up, EOI option open on website)	March 2023	S&L UK/EEF
Final ethics approval	March 2023	IES
Updated proposal and evaluation budget	7 March 2023	IES
Grants committee (sign-off of the evaluation budget)	31 March 2023	EEF/IES
Recruitment of settings	April – September 2023	S&L UK
Screening to identify the population of interest (WellComm)	October – November 2023 (5 weeks)	IES and Qa
Baseline assessment for those identified as eligible as a result of the screening (PLS-5)	November – December 2023	IES and Elklan
Practice as usual survey	October – November 2023	IES
Randomisation	December 2023	IES
Informing settings of their randomisation allocation	January 2024	S&L UK
Early Talk Boost training start (for those allocated to intervention group) and evaluation observations of training	January 2024	S&L UK / IES
Implementation in settings	January – March 2024	S&L UK
Online observations of support sessions	February – March 2024	IES
Evaluation case studies, including interviews with parents and interviews with settings/practitioners that have dropped out	March – April 2024	IES
Interviews with delivery team	April – May 2024	IES
Collection of costs data from settings (TBC)	April 2024	IES
Endline staff survey(s)	End of April – May 2024	IES

Endline child assessment (FOCUS-34, PLS-5)	End of April– May 2024	IES and Elklan (PLS-5)
Second IDEA workshop and sharing of relevant fidelity/delivery data	June 2024	IES / S&L UK
Report first draft due	October 2024	IES
Practitioner training starts for waitlist control group	September – October 2024	S&L UK
Report published	Spring 2025	EEF

References

Dawson, A., Huxley, C., & Garner, O. (2022) *Coaching Early Conversation, Interaction and Language (CECIL) Evaluation: Implementation and Process Evaluation*, IES Report 574, Institute for Employment Studies. https://www.suttontrust.com/wp-content/uploads/2022/03/CECIL_IPEreport-FINAL-02.03.22.pdf

EEF (2022) *Implementation and process evaluation guidance for EEF evaluations, August 2022*. Education Endowment Foundation. https://d2tic4wvo1iusb.cloudfront.net/documents/evaluation/evaluation-design/EEF-IPE-Guidance-August-2022.pdf?v=1660741017

González, S., & Bonal, X. (2021) 'COVID-19 school closures and cumulative disadvantage: Assessing the learning gap in formal, informal and non-formal education', *European Journal of Education*, 56, 607-622. https://doi.org/10.1111/ejed.12476

Hsiao, Y., Qi, C. H., Hoy, R., Dale, P. S., Stump, G. S., Davison, M. D., & Xia, Y. (2021) 'Hierarchy and Reliability of the Preschool Language Scales–Fifth Edition: Mokken Scale Analysis. *Journal of Speech, Language, and Hearing Research*', 64(10), 3983-3994. https://doi.org/10.1044/2021_JSLHR-21-00043https://doi.org/10.1044/2021_JSLHR-21-00043

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R & Kerr, K. (2016) Implementation and process evaluation (IPE) for interventions in education settings: An introductory handbook. The Education Endowment Foundation, London. https://d2tic4wvo1iusb.cloudfront.net/documents/evaluation/evaluation-design/IPE_Handbook.pdf?v=1684351917

Law, J., Charlton, J., and Asmussen, K. (2017) 'Language as a child wellbeing indicator'. *Early Intervention Foundation,* [online] Available at: <https://www.eif.org.uk/report/language-as-a-child-wellbeing-indicator> [Accessed 26 April 2023].

Marshall, J., Adelman, A., Kesten, S. M., Natale, R. A., Elbaum, B. (2017) 'Parents' experiences navigating intervention systems for young children with mild language delays', *Journal of Early Intervention*, 39(3), 180-198. https://doi.org/10.1177/1053815117704958

Oddson, B., Thomas-Stonell, N., Robertson, B., & Rosenbaum, P. (2019) 'Validity of a streamlined version of the Focus on the Outcomes of Communications Under Six: Process and outcome', *Child: Care, Health and Development,* 45(4), 600-605. **https://doi.org/10.1111/cch.12669**

Oughton, R. (2022) Sample size calculations.

Reeves, L., Hartshorne, M., Black, R., Atkinson, J., Baxter, A., & Pring, T. (2018) 'Early talk boost: A targeted intervention for three year old children with delayed language development', *Child Language Teaching and Therapy*, 34(1), 53-62.

Speech and Language UK. (2015) 'Early Talk Boost Evaluation Report 2015', *Speech and*

Language UK, [online] Available at <https://speechandlanguage.org.uk/media/4297/early-talk-boost-evaluation2015.pdf>

Thomas-Stonell, N.L., Oddson, B., Robertson, B. and Rosenbaum, P.L. (2010) 'Development of the FOCUS (Focus on the Outcomes of Communication Under Six), a communication outcome measure for preschool children', Developmental Medicine & Child Neurology, 52, 47-53. https://doi.org/10.1111/j.1469-8749.2009.03410.x

Zimmerman, I. L., Steiner, V, G., & Pond, E. (2011) Preschool Language Scales- Fifth Edition UK (PLS-5 UK). San Antonio, TX: Pearson

Appendix A: Assumed causal mechanisms

Appendix Table 1: Assumed causal mechanisms, with the evidence strength assessed by Speech and Language UK.

#	Assumption / Causal Mechanism	Where in the ToC does the assumption appl		Evidence	Evidence Strength
1	oral language interventions improve outcomes for children	Short-term outcome	Long-term outcome	West, G., Snowling, M.J., Lervåg, A., Buchanan-Worster, E., Duta, M., Hall, A., McLachlan, H. and Hulme, C. (2021), Early	Green
2	children's communication and language skills are improved through adult use of contingent language behaviours		Short-term outcome	McGillion, M., Pine, J.M., Herbert, J.S. and Matthews, D. (2017), A randomised controlled trial to test the effect of promoting	Green
3	EYPs use strategies they learn through training and intervention	Short-term outcome	Long-term outcome	Self reported fidelity check on I CAN's online tracker	Amber/Red
4	children acquire active listening skills through positive real time reinforcements as scripted in the intervention		Short-term outcome	Gilmore J and Vance M (2007) Teacher ratings of children's listening difficulties. Child Language Teaching and Therapy 23:	Amber/Green
5	children are exposed to the key elements of vocab acquistion; semantic and phonological cues and use these		Short-term outcome	Justice L, Meier J, and Walpole S (2005) Learning new words from storybooks: An efficacy study with at risk	Green
6	children extend their utterance length because of modelling using SVO and SVA sentence structures		Long-term outcome	Girolametto, L. and Weitzman, E., 2002. Responsiveness of child care providers in interactions with toddlers and	Green
7	EYPs select the children for whom the intervention is designed		Input/Activity	Prior to ETB intervention 51% children rated amber (40% red and 9% green); post intervention 59% green (32% amber and	Amber/Red
8	the training gives EYPs sufficient understanding of the intervention and how to run it		Output	88% of respondents to I CAN's training feedback survey said they feel the training has prepared them to deliver the	Amber/Red
9	children make accelerated progress with their communication and language after ETB intervention		Long-term outcome	Reeves L, Hartshorne M, Black R, Atkinson J, Baxter A, Pring T. Early talk boost: A targeted intervention for three year old	Amber/Green
10	parents sharing Jake and Tizzy books with their children will impact on their communication and language skills	Short-term outcome	Long-term outcome	Colmar S (2014) A parent-based book-reading intervention for disadvantaged children with language difficulties. Child	Amber/Green