Project Title: Independent evaluation of the Early Years Conversation Project (EYCP): A two-armed cluster randomised waitlist-controlled trial
Evaluation Protocol

Evaluator (institution): Durham University
Principal investigator(s): Xiaofei Qi and Victoria Menzies

Evaluation summary

Project title	Independent evaluation of the Early Years Conversation Project (EYCP): A two-armed cluster randomised waitlist-controlled trial
Developer *(Institution)*	East London Research School
Evaluator *(Institution)*	Durham University
Principal investigator(s)	Xiaofei Qi and Victoria Menzies
Protocol author(s)	Xiaofei Qi, Victoria Menzies, Nadia Siddiqui, Nashwa Ismail, and Rachel Oughton
Trial design	Two-armed cluster randomised waitlist-controlled trial with random allocation at the nursery level
Trial type	Efficacy
Pupil age range and Key stage	2 to 3 year olds; Early Years
Number of schools *(at design stage)*	112
Number of pupils *(at design stage)*	1120
Primary outcome measure and source	Receptive and expressive vocabulary measured with the British Picture Vocabulary Scale-Third Edition (BPVS-III) (researcher administered), the Early Language Identification Measure (ELIM) (parent/carer administered) and the MacArthur-Bates Communicative Development Inventories-III (MB-CDIs-III) (parent/carer administered)
Secondary outcome measure and source	1) Children's language and communication skills measured with Wellcomm Early Years: The complete speech and language toolkit (henceforth referred to as Wellcomm Early Years toolkit) (researcher administered) 2) Practitioner's pedagogical knowledge related to early language development measured with the Observing Language Pedagogy (OLP) tool (practitioner self-report)

Protocol version history

Version	Date	Reason for revision
1.0 [*original*]	19/09/2023	N/A

Contents

Evaluation summary 1
Protocol version history 2
Study rationale and background 5
Intervention 6
Impact evaluation design 12
Implementation and process evaluation design 27
Cost evaluation 36
Ethics and registration 37
Data protection 37
Personnel 38
Risks 39
Timeline 40
References 42
Appendix A: EYCP Theory of Change 46
Appendix B: Contextual assumptions log 47
Appendix C: Assumption/causal mechanism 49
Appendix D: Changes since the previous evaluation 51
Appendix E: EEF-SPH recruitment strategy template 52
Appendix F: Sample size calculations 53

Tables

Table 1: TIDieR checklist 6
Table 2: Trial design 13
Table 3: Sample size calculations (average 10 children per setting; no attrition) 20
Table 4: Sample size calculations (minimum 6 children per setting; no attrition) 20
Table 5: Sample size calculations (average 10 children per setting; school-level attrition: 5%; Child-level attrition: 10%) 22
Table 6: EYCP Compliance Criteria 25

Table 7**:** IPE design, methods of data collection and analysis overview 33

Table 8: Risks 39

Table 9: Timeline 40

Figures

Figure 1: Logic model of EYCP Efficacy Trial 11

Study rationale and background

Early language ability is one of the strongest predictors of later development through school and in life (Burchinal et al. 2016; Hoff, 2013; Pace et al., 2017). Language is essential for thinking, social and emotional wellbeing and learning in all areas of development. The significance of early language development and communication skills has been widely recognised in the early years foundation stage (EYFS) (Department for Education, 2023). Language and communication in the early years are fundamental skills for future learning and can predict language outcomes and reading skills throughout primary school; the persistence of low language skills is linked with lower academic achievement (Snow, Burns and Griffin, 1998, Eadie et al, 2021; Sénéchal, Ouellette and Rodney, 2006).

Influences such as the home learning environment and the preschool learning environment have been shown to play essential roles in children's development at early ages (Lehrl, Evangelou, and Sammons, 2020; Melhuish et al., 2008; Sylva et al., 2004). Existing evidence shows that providing children with the experience of a varied and rich language and literacy exposure at home and in preschool provision, such as sharing books, conversations, and using child-directed speech, promotes children's language and literacy development (Mol and Neuman, 2014; Hoff, 2013; Law, Charlton and Asmussen, 2017).

It is well documented that the language skills of preschool children differ substantially and that these differences are highly predictive of their later academic success and achievements. The prevalence of language delays and difficulties in early years is of great concern (Public Health England,2020) and it is argued that cost-effective, evidence-based training and interventions that promote the most effective types of language-boosting interactions between children and those caring for them (parents and early years practitioners) are needed to ensure that all children have the best possible chance of reaching their full potential (Law et al., 2017; Wooles, Swann and Hoskison, 2018).

Communication and language approaches that emphasise the importance of spoken language and verbal interaction for young children have shown promising results in the EEF Early Years Toolkit (EEF, n.d.). Many children from disadvantaged backgrounds have poorer communication, language & vocabulary (Law et al., 2017; Wooles, Swann and Hoskison, 2018). There is existing evidence of language gaps for disadvantaged pupils at the beginning of school; children from low-income backgrounds consistently perform below their more advantaged peers on standardized measures of language ability, setting long-term trajectories that translate into gaps in academic achievement (Pace et al., 2017; Schwab and Lew-Williams, 2016). According to statistics reports on government-funded early years provision for children under 5 years of age in England, most disadvantaged children are in private nursery settings (PVIs). For example, 625,434 children under age 4 are registered for 15-hour entitlement with PVI providers in 2023, and only 298,903 children are registered with state-funded schools (see Education provision: children under 5 years of age; Gov.UK., 2023a). However, staff in PVIs have lower levels of qualifications and less access to professional development (Haux et al., 2020), including professional development (PD) focused on communication and language, when compared to staff in schools.

A small number of studies that have taken place in nurseries with a higher proportion of children experiencing socio-economic disadvantage tended to have above average effects suggesting that this is likely to be a beneficial approach for this group (Law et al., 2017). Hence targeted language and communication support in early years may be a promising approach to narrow these inequalities and learning gap before school age. The Early Years Conversation

Project (EYCP, formerly piloted as Manor Park Talks), developed and delivered by East London Research School (ELRS), based at Sheringham Nursery School and Children's Centre, is a PD programme designed to support early years/nursery staff with developing understanding and practical strategies to enhance children's communication and language skills. It is backed up by a strong body of research evidence which consistently shows that communication and language approaches benefit young children's learning (EEF, n.d),and aims to promote conversational responsiveness by training early years professionals to implement and embed the 'ShREC approach' (Share attention, Respond, Expand and Conversation) in everyday practice.

As part of the Department for Education's (DfE) Early Years Recovery Programme, A Brighter Start Stronger Practice Hub and the Education Endowment Foundation (EEF) are working together to fund early years settings with access to the EYCP. An evaluation team (ET) from the University of Durham has been appointed to evaluate the impact of EYCP on young children's communication and language development.

An EEF-funded pilot evaluation of the programme was conducted by NatCen, across 12 early Years settings and 6 schools in Newham (Husain, Morris, Basi, and Bristow., 2020). Findings from the pilot demonstrated evidence of promise for improving children's communication and language outcomes, as hypothesized in the programme's theory of change (ToC). The pilot results also indicated that the strategies were becoming embedded in daily practice. Given this context and existing evidence, we have developed an efficacy trial that integrates a randomised control trial to evaluate the impact of the EYCP on 2 to 3 year old nursery children's language and communication skills with an implementation and process evaluation (IPE) to explore the causal pathways of the ToC and identify elements of effective delivery. (See Table 2 for evaluation design). The efficacy trial will allow us to rigorously investigate the potential impact of the programme under best possible conditions. It also adds to the evidence base on the benefits of a language programme that supports staff working with 2 to 3 year olds, an area where there is currently an evidence gap.

Intervention

Table 1 describes the EYCP in detail using the Template for Intervention Description and Replication (TIDieR) checklist.

Table 1: TIDieR checklist[1]

TIDieR Item	Description
Brief name	Early Years Conversation Project (EYCP)
Why: Rationale, theory and/or goal of essential elements of the intervention	**Background** • A key determinant of language development is the amount and quality of language to which a young child is exposed. • Many children from disadvantaged backgrounds have poorer communication, language & vocabulary. • Most disadvantaged children are in private, voluntary or independent (PVI) nursery settings. • Staff in PVIs have lower levels of qualifications & less access to professional development, including PD focused on

[1] The TIDieR checklist was written by Melissa Prendergast and Fliss James from the delivery team, and with revisions from the ET.

	communication and language, when compared to staff in schools.
	
	EYCP overview
	The EYCP focuses on conversational responsiveness – using adult-child interactions to increase the amount a child says and the complexity of their conversation. The aim of the programme is to improve the skills and expertise of the early years workforce to better implement techniques that promote conversational responsiveness, to support children's language development.
	• PD for practitioners in early years settings to use and understand effective strategies to help children develop their communication and language, based around the EEF's 'ShREC approach' (James, 2022).
	• The 'ShREC approach' incorporates four evidence-informed strategies:
	  ○ Share attention
	  ○ Respond
	  ○ Expand
	  ○ Conversation
	• This approach is intended to promote conversational responsiveness throughout the child's day in nursery.
	• Key features of the approach are multi-turn conversations and decontextualised talk with a focus on vocabulary.
	• Interactive book reading is a key vehicle for this approach.

Who: Recipients of the intervention	**Nursery** A PVI nursery, a state-funded school-based nursery (SN) or a maintained nursery offering Free Early Education and Childcare (FEE) for 2-year-olds from the chosen local authorities in London (see 'Where: Location of the intervention' for detail). **Practitioner** Two members of staff from each setting will take part in the EYCP PD: 1) The setting manager/ Early Years Foundation Stage (EYFS) lead, and 2) The EYCP champion, a room lead/nursery teacher working with 2 to 3 year olds **Children** The EYCP programme was developed to improve the language and communication provision and attainment in language for all children in participating nurseries. Nonetheless, 2 to 3 year old children in the EYCP champion's classroom will be the direct recipients of the programme
What: Physical or informational materials used in the intervention	• **Professional development** through a mix of face to face and virtual training. This will involve 1.5 days face to face and 6 hours of online recorded modules. The recordings will be held on an online platform (i.e., Padlet). • **Specialist mentoring** through a mix of face to face and virtual hub meetings and setting support visits from mentors. • **Access to a dedicated online knowledgebase.** The EYCP online platform will extend PD learning by providing video exemplification, background materials, and information and ideas to support families at home. • **Physical handbook** with programme timetable and photocopiable resources.

What: Procedures, activities and/or processes used in the intervention	• **PD training** with a mix of face to face and virtual sessions. 1.5 days face to face and 6 hours of online recorded modules.
	  ○ The EYCP Champion (room leader or nursery teacher working directly with children aged 2 to 3 years old) will receive all the training.
	  ○ The setting manager or EYFS leader will join them for specific sessions.
	• **Specialist mentoring for the EYCP champions** through a mix of face-to-face and virtual hub meetings.
	  ○ Mentors will run 9 hub meetings across the year with a minimum of 5 support visits from mentors.
	  ○ Mentors will mentor a specified number of settings during the intervention period. If possible, the same group of mentors (n=8) will stay in the EYCP programme over three years: in the 2023-2024 school year (Cohort 1), 8 mentors will work in pairs, each pair to support 7 settings; in the 2024-2025 school year (Cohort 2) each mentor will independently support 7 settings; and in the final year (2025-2026) the 8 mentors will work in pairs again to support 7 settings.
	  ○ At the beginning and end of the programme an ITERS-3 (Infant/Toddler Environment Rating Scale) (Harms et al., 2017) observation of the setting's environment is completed by the mentor to provide formative feedback on progress towards providing a language-rich environment.
	  ○ Hub meetings will be used to build knowledge, revisit learning from the online modules and embed understanding. Mentors will use video to help participants tune in to what the strategies look like, how successfully they use them and the impact on the children. These meetings will provide needs-based support for the EYCP Champion to implement the strategies, overcome barriers, and cascade strategies to their wider team.
	• **EYCP champions to implement and embed the ShREC strategies** - Share attention, Respond, Expand and Conversation - with young children in everyday practice.

Who: Intervention providers/implementers	**PD providers**
	• The PD training is delivered by Fliss James and Melissa Prendergast (referred to here as the ELRS team).
	• A team of mentors (n=8) provide the in-setting support and run hub meetings.
	  ○ Mentors are trained by the ELRS team prior to the intervention starting. This includes a session on essential evidence, strategies, and features of effective PD.
	  ○ Mentors are trained in the use of ITERS-3 and then moderated. Mentors attend all the PD training alongside participating practitioners. They all attend regular meetings as a team, with ELRS.
	
	Intervention implementers
	• The EYCP champion implements the strategies of EYCP in their setting with all children in their classroom every day. They cascade the training to their team through an informal buddy system.
	• The setting manager engages with PD on programme implementation and supports the EYCP champion in implementing the strategies.

How: Mode of delivery	The EYCP includes a mix of online and face to face PD training and mentoring sessions. PD sessions, hub meetings and support visits run in an alternating cycle.

Where: Location of the intervention	The EYCP is available to nursery settings from the following local authorities: Barking and Dagenham, Barnet, Brent, Camden, City of London, Enfield, Greenwich, Hackney, Hammersmith and Fulham, Haringey, Havering, Islington, Kensington and Chelsea, Lambeth, Lewisham, Newham, Redbridge, Southwark, Tower Hamlets, Waltham Forest & Westminster.
When and how much: Duration and dosage of the intervention	The EYCP is 8 months in length, and it will be delivered in the 2023-2024 school year (Cohort 1) and then repeated in the 2024-2025 school year (Cohort 2 and Cohort 1 waitlist delivery), and the following 2024-2025 school year (Cohort 2 waitlist delivery). • Setting managers will attend 1.5 days face to face training and 1 recorded online training session (1 hour; focus on content & behaviour change). • EYCP Champions will attend 1.5 days face to face training, all 6 recorded online training sessions, and 9 hub meetings. They will receive one mentor support visit per half term, with extra visits if mentors feel they are needed. • The EYCP champion is expected to use the strategies of EYCP every day and with every child in their classroom.

Tailoring: Adaptation of the intervention	• The PD training and mentor support is manualised but there is the opportunity to provide bespoke support to participants, based on need.
	• All aspects of the intervention are evaluated and monitored; this allows the team to pick up where bespoke support might be needed.
	  ○ The PD sessions delivered by the ELRS team are evaluated and any misconceptions or gaps in knowledge are responded to by meeting with mentors and supporting them to address these in hub meetings and support visits.
	  ○ Mentors reflect and evaluate on their work with PD recipients. These self-evaluations are shared with the ELRS team.
	• The two teams - the ELRS team and the team of mentors - meet regularly and work together to deliver bespoke support where needed.
	• The ELRS team have oversight of this tailoring which allows for fidelity/adherence control and checks.

How well (planned): Strategies to maximise effective implementation	• The ELRS team have delivered 3 iterations of this programme prior to EYCP. This has allowed for a long explore and prepare stage. • The PD programme is manualised with careful adherence to the EEF's guidance report on professional development. The PD programme has a balanced design with mechanisms in all four areas: build knowledge, motivate staff, develop teaching techniques, and embed practice. This structure is adhered to by the ELRS team and the mentors. • Mentors use Padlet to access the training materials available to participants. They also have a dedicated mentor area that contains mentor-specific training. It also holds the mentor reflection and evaluation documents, which are carefully monitored by the ELRS team. The highly trained and supported mentors work closely with the participants in their hub to ensure that the strategies in EYCP are implemented successfully in their setting. • The ELRS team supports participants to scale and sustain the strategies in EYCP. This is done with a PD session on implementation and support from mentors and peers.

A detailed ToC for the EYCP programme was developed by the delivery team (DT), ET and EEF in the set-up meetings (see Appendix A) along with the associated contextual and causal assumption logs in Appendix B and Appendix C. Changes since the previous evaluation of the Manor Park Talks can be found in Appendix D.

A more specific logic model was developed for this efficacy trial (see Figure 1). The causal mechanism of the logic model is that providing PD for practitioners in early years settings to use and understand effective language and communication strategies based around the 'ShREC approach' (James, 2022) can improve children's vocabulary, language and communication skills. The targeted population are 2 to 3 year olds in PVI, maintained nurseries, and SN settings offering 15 hours of FEE to 2 year olds; setting managers/EYFS lead and EYCP champions. The detailed EYCP intervention inputs/activities, outputs and expected outcomes for targeted population can be seen in Figure 1.

Figure 1: Logic model of EYCP Efficacy Trial

Causal mechanism: increasing early years practitioners' use and understanding of effective language and communication strategies based around the 'ShREC approach' will enhance practitioner's language-related pedagogical knowledge and improve children's vocabulary, language, and communication skills.

Inputs/Activities	Outputs	Outcomes
Setting managers/EYFS leads: • PD training through a mix of face to face (1.5 days) and virtual session (1 session) (Core activity) • Access to EYCP programme handbooks dedicated online knowledgebase (Optional activity) • Engage with online networking activity and programme support (Optional activity) **EYCP champions:** • PD training through a mix of face to face (1.5 days) and virtual sessions (6 pre-recorded sessions) (Core activity) • Specialist mentoring through a mix of face-to-face and virtual 'hub' meetings (9 hub meetings, minimum 5 in-setting visits and 2 IERS-3 audits from mentors) (Core activity) • Access to EYCP handbooks and dedicated online knowledgebase (Core activity) • Strategies cascaded from mentor to EYCP champion via modelling, videoing & reflection (Core activity) • Engage with online networking activity and programme support (Optional activity) **Children:** experience a richer language environment, with higher quality early years teaching involving regular multi-turn conversations through the use of the ShREC approach, including interactive book reading.	**Setting managers/EYFS leads:** • Engage with PD training & robust evidence about communication and language • Support EYCP Champion behavior change and EYCP implementation **EYCP champions:** • Engage with robust evidence about communication and language • Engage with PD, mentoring & in-setting support visits from mentors. • Engage with online networking opportunities and programme support. • Engage with children using the 'ShREC Approach' • Cascades strategies to their wider team **Children:** experience practitioners sharing attention with them; engage with practitioners in sustained contingently responsive conversation; hear increasingly complex language	**Primary outcome: (Impact)** • Improvement on children's receptive & expressive vocabulary attainment **Secondary outcomes: (impact)** • Improvement on children's language and communication skills • Improvement on practitioners' language-related pedagogical knowledge **Exploratory outcomes: (process and impact)** • Improvement on practitioner confidence in language teaching • Observed changes of language provision learning environment

Potential mediators/moderators: Children: FEE and EAL status at age 2, Baseline performance on vocabulary attainment; Practitioner: teaching qualification and teaching experience

Impact evaluation design

Research questions

In the impact evaluation, our primary research question is:

1. What is the impact of the EYCP on the receptive and expressive vocabulary [measured by the BPVS-III, the ELIM and the MB-CDIs-III] of 2 to 3 year olds in intervention nurseries receiving the EYCP in comparison to those children in the waitlist-control nurseries receiving business as usual? *(Primary outcome)*

Our secondary impact evaluation research questions are:

2. What is the impact of the EYCP on the language and communication skills [measured by the Wellcomm Early Years toolkit] of 2 to 3 year olds in intervention nurseries receiving the EYCP in comparison to those children in the waitlist-control nurseries receiving business as usual? *(Secondary outcome)*

3. What is the impact of the EYCP on practitioner's language-related pedagogical knowledge [measured by the OLP tool] in intervention nurseries receiving the EYCP in comparison to those practitioners in the waitlist-control nurseries receiving business as usual? *(Secondary outcome)*

4. What is the impact of the EYCP on the receptive and expressive vocabulary of children who are *eligible for free education and childcare (FEE) at age 2* in intervention nurseries receiving the EYCP in comparison to that subgroup of children in the waitlist-control nurseries receiving business as usual? *(Primary outcome: FEE subgroup)*

5. What is the impact of the EYCP on the receptive and expressive vocabulary of children with English as an additional language (EAL) in intervention nursery settings receiving the EYCP in comparison to that subgroup of children in the waitlist-control nurseries receiving business as usual? *(Primary outcome: EAL subgroup)*

6. What is the impact of the EYCP on the receptive and expressive vocabulary of children with least/best baseline performance [measured by the BPVS-III and ELIM] in intervention nurseries receiving the EYCP in comparison to that subgroup of children in the waitlist-control nurseries receiving business as usual? *(Primary outcome: best/least performance subgroup)*

Design

We have designed a two-armed cluster randomised waitlist-controlled efficacy trial with randomisation at the nursery level to evaluate the impact of the EYCP on progress in the vocabulary, language and communication skills of children aged 2 to 3 years old in nursery settings (See Table 2 for trial design). A waitlist-control design is chosen given the EYCP is co-funded with the Department for Education (DfE)'s Stronger Practice Hub (SPH) initiative and consequently there is an interest in all settings accessing the intervention. The SPH initiative is part of the early years education recovery support package to provide advice, share good practice, and offer evidence-based professional development for early years practitioners (DfE, 2022).

The trial will be run over three years, with the intervention being delivered to two different cohorts. Cohort 1 settings will be recruited and randomised into intervention and control groups in 2023, with intervention settings receiving the EYCP during the 2023/2024 school year and waitlist-control settings receiving it during the 2024/2025 school year. Cohort 2 settings will be recruited and randomised into intervention and control groups in 2024, with intervention settings receiving the EYCP during the 2024/2025 school year and waitlist-control settings receiving it during the 2025/2026 school year.

Table 2: Trial design

Trial design, including number of arms		Two-armed cluster randomised waitlist-controlled trial
Unit of randomisation		Nursery
Minimisation Factors		• Nursery type (PVIs, maintained and SN settings) • Nursery size with 2 to 3 year old cohort (<median number of children; ≥ median number of children) • Geographic location (Local Authorities as proxy for 'mentor area')
Primary outcome	**Variable**	Receptive and expressive vocabulary
	Measure (instrument, scale, source)	• BPVS-III (0-168, raw score), researcher administered, GL Assessment • MB-CDIs-III (0-100, raw score), parent/carer administered, Brookes publishing
Secondary outcome(s)	**Variable(s)**	1) Children's language and communication skills 2) Practitioner's pedagogical knowledge of language learning and development
	Measure(s) (instrument, scale, source)	1) Wellcomm Early Years toolkit (raw score), researcher administered, GL assessment 2) OLP tool (raw score), practitioner self-report
Baseline for primary outcome	**Variable**	Receptive and expressive vocabulary
	Measure (instrument, scale, source)	• BPVS-III (0-168, raw score), researcher administered, GL assessment • ELIM (0-50, raw score), parent/carer administered, Public Health England
Baseline for secondary outcome	**Variable**	N/A
	Measure (instrument, scale, source)	N/A

Participant selection

Nurseries

As agreed in the IDEA workshop and set-up meetings with the DT from ELRS and EEF, we will recruit 112 nurseries over two years (56 in 2023 and 56 in 2024) that:

- Are located in London[2] (specifically Barking and Dagenham, Barnet, Brent, Camden, City of London, Enfield, Greenwich, Hackney, Hammersmith and Fulham, Haringey, Havering, Islington, Kensington and Chelsea, Lambeth, Lewisham, Newham, Redbridge, Southwark, Tower Hamlets, Waltham Forest & Westminster);
- Are a PVI, a maintained or a SN nursery offering FEE for 2 year olds[3];
- Have a minimum of 8 children aged 2 years old at the start of 2023-2024 or 2024-2025 school year, who will attend the setting for a minimum of 15 hours a week during the relevant school year;
- Will be able to commit to releasing one practitioner who works directly with 2 to 3 year olds as the EYCP Champion for PD training sessions (including 1.5 days face to face and 6 hours of recorded modules), 9 hub meetings and a minimum of 5 support visits from mentors;
- Will be able to commit to releasing the setting manager or the EYFS lead for 3 training sessions, in addition to supporting the EYCP Champion throughout the project; and
- Have not participated in Manor Park Talks or Newham Communication Project; and
- Are not taking part in the DfE's Early Years Professional Development Programme or any other Early Years trials funded by the EEF or similar funder.

While the EYCP programme was originally developed to support practitioners in PVI settings, in this efficacy trial we will include maintained nurseries and SN settings who are providing 15 hours of FEE to 2 year olds in selected local authorities. The reason is twofold: 1) to represent the main types of EYFS providers for 2 to 3 year olds, and 2) to reduce the overall attrition rate which is particularly higher in PVI settings due to reasons such as staff turnover. We are aiming to recruit 56 nurseries in each cohort (28 in the intervention arm and 28 in the waitlist-control arm), of which at least 12 (21%) are SN or maintained school nurseries and 44 (79%) are PVI settings, which represents the national split of early years provider by type providing 15 hours FEE for 2-year-olds in 2023 (see Education provision: children under 5 years of age; Gov.UK, 2023a).

The DT is leading the recruitment process, supported by the ET. A detailed recruitment strategy was developed by the DT, in line with the EEF-SPH recruitment guidance (see Appendix E for the recruitment strategy template). Some key recruitment activities and methods include:

- An EYCP logo is included on all ELRS's social media outputs and documents.
- Web page on the SPH section of the DT's nursery website: EYCP - Sheringham Nursery School (sheringham-nur.org.uk). The EYCP webpage will summarise the project, including eligibility, links to expression of interest (EOI) forms and programme sign-up.
- The DT approach their contacts in the Greater London Authority (GLA) and Local Authorities (LAs) of the selected areas. These will be heads of Early Years (EY) and officers in charge of EY improvement and CPD.
- The DT do short elevator pitches of the project at events and meetings with system leaders (GLA, nursery heads, etc)

[2] There is the possibility that recruitment will be expanded to another regional hub outside of London for Cohort 2 (2024-2025 school year).
[3] Considering the upcoming childcare policy reform in England (see Childcare choices; Gov.UK., 2023b) whereby all working families with 2 year olds will receive 15 hours of FEE from April 2024, the Income Deprivation Affecting Children Index (IDACI) rank of the setting's postcode will be used as a proxy for disadvantage instead of 'offering FEE for 2 year olds' for Cohort 2 (2024-2025 school year).

- Using the databases shared by LAs and the DfE, the DT make contact with settings, promoting the project through flyers, emails and phone calls.
- The DT will promote the project at events in the specified areas (liaise with LAs to ensure they are invited to managers' forums, and training events and have a space in newsletters). They will work with the GLA on joint events and use this as an opportunity to promote the EYCP.
- The EYCP will be promoted through A Brighter Start SPH and DT's social media accounts.
- The EEF will promote the EYCP through their social media, website, newsletters, blogs, DfE and early years contacts, as well as early years publications.

Settings who are interested in participating and completing the EOI will be checked by the DT against the specified setting eligibility criteria. Once eligibility has been confirmed, nurseries will be recruited on a first come first served basis, with location and nursery type taken into consideration. Eligible settings will be provided with a Memorandum of Understanding (MoU) by the DT, which provides full information relating to the setting's involvement within the trial, alongside a Data Sharing Agreement (DSA) between the nursery, the DT (from ELRS) and the ET (from Durham University). The DT will send out and follow up on the return of the MoU and DSA and share them with the ET. After 56 nurseries have been recruited, eligible settings who are returning trial paperwork will be placed on a reserve list; they will be called upon should a recruited nursery drop out before the randomisation process.

All participating nurseries will receive the EYCP programme for free: settings assigned to the intervention arm will receive the programme during the intervention period (Cohort 1: 2023-2024 school year; Cohort 2: 2024-2025 school year), whereas waitlist-control settings will receive the programme in the following school year (Cohort 1: 2024-2025 school year; Cohort 2: 2025-2026 school year). In addition, there will be financial incentives for all participating nurseries in this efficacy trial (i.e., both intervention and waitlist-control settings), with £150 provided at the start of the trial (after baseline assessment) and £250 at the end of the trial (after post-intervention assessment). Moreover, there will be 50% of cover costs for staff attendance at EYCP training in all participating settings. The DT will manage all setting-level financial incentives for this trial.

Nursery practitioners

When signing the setting MoU, participating settings are required to nominate two practitioners to take part in the EYCP PD: 1) the setting manager/ EYFS lead, and 2) the EYCP champion, a room lead/nursery teacher working with 2 to 3 year olds. The EYCP champion information will be confirmed by the setting manager/EYFS lead before the baseline assessment happens.

In addition to setting-level incentives, practitioners (the EYCP champions) who are completing the OLP tool at the end of the EYCP intervention or participating in the IPE interviews/focus groups will be provided with a £15 high-street voucher or local grocery store voucher as recognition for participating in the research. The ET will manage all practitioner-level financial incentives for this trial.

Children

Child eligibility criteria:

- Children aged 2 years old at the start of nursery in September 2023 (Cohort 1) or September 2024 (Cohort 2);

- Children who attend nursery for a minimum of 15 hours per week during term time in the 2023-2024 school year (Cohort 1) or the 2024-2025 school year (Cohort 2);
- Children in the class of the nominated EYCP champion;
- Children whose parents/carers anticipate that they will remain at the nursery (i.e., they do not foresee they will leave the nursery) for the duration of the trial (until June); and
- Children who complete the baseline assessment.

Children are not eligible if they have significant special educational needs and disabilities (SEND) which would prevent them from accessing the assessment and/or would be distressed through completing the assessment.

When signing the MoU and DSA, settings will be asked to provide the estimated number of 2-year-old children starting nursery in September as well as an estimated number of children who are eligible for 15-hour FEE at age 2[4]. This will allow the ET to have an estimation of the median cohort size of 2 to 3 year olds in participating settings.

After returning the MoU and DSA, participating settings will be provided with digital and/or paper information sheets and privacy notices for parents/carers with instructions on opting into the EYCP efficacy trial. Settings are encouraged to begin obtaining parent/carer consent over the summer before the start of the school year, contacting parents/carers of children who are on the setting's pre-registration lists via email. The DT will produce a video for parents/carers explaining the EYCP programme and evaluation activities.

In early September, all settings will provide the ET with the number of eligible children meeting the first three eligibility criteria and support the ET in distributing the EYCP information sheets, privacy notices and opt-in consent forms to parents/carers of eligible children. In this trial, we are aiming to recruit a minimum of 8 and a maximum of 15 eligible 2 to 3 year old children per setting. For participating settings with more than 15 eligible children, 15 children will be randomly selected and recruited by the ET to take part in the baseline assessment. Remaining eligible children well be placed on a reserve list; they will be called upon should a recruited child drop out during baseline assessment and/or before setting randomisation.

Parents/carers of participating children will be asked to complete two questionnaires, one at the beginning of the intervention in September/October, one at the end of the intervention in June/July; a £15 high street or local grocery store voucher will be provided each time for completing the questionnaire as a recognition for participating in the research. The ET will manage all family-level financial incentives for this trial.

Outcome measures

Baseline measures

At baseline, we will use the researcher-administered BPVS-III (Dunn et al, 2009) together with the parent/carer-reported ELIM (Law et al., 2020; Public Health England, 2020) to capture children's receptive and expressive vocabulary attainment before the EYCP intervention.

BPVS-III

[4] Considering the upcoming policy changes to childcare support in England (see Childcare choices; Gov.UK., 2023b) whereby all working families with 2 year olds will receive 15 hours of FEE from April 2024, the Income Deprivation Affecting Children Index (IDACI) rank of the child's postcode (collected from their parent/carer) will be used as a proxy for disadvantage instead of FEE eligibility for Cohort 2 (2024-2025 school year).

The BPVS-III is a standardised vocabulary measure based on the Peabody Picture Vocabulary Test (PPVT: Dunn and Dunn, 1981). The scale is designed to assess receptive (hearing) vocabulary with children aged 3 to 16 years old and can be used as a baseline measure for children starting nursery or school, as well as a benchmarking and progress-checking assessment. It is used widely by speech and language therapists, educational psychologists, and additional support for learning teachers because vocabulary is considered to be an important predictor of other language skills (Dunn et al., 2009). The testing involves showing the child a series of images (four on one page) and asking the child to point to the picture that matches a pre-specified word the assessor says to the child.

Given no reading or verbal response from children is needed for the BPVS-III assessment, it is considered suitable for young children and those for whom English is an additional language. It takes a relatively brief time (between 5 to 8 minutes) to complete, enabling children to remain focused during the assessment period. BPVS-III has been standardised in the UK for children aged 3 to 16 years; it has also been used in studies (i.e., ParentChild+; Tracey et al., 2019; 2022) with 2 to 3 year olds and it was regarded as a suitable measure in that there was variation in response and no floor effects were detected.

The BPVS-III (raw scores; 0-168) will form part of the baseline assessment outcome measuring the receptive vocabulary of children. It will be administered one-to-one by an independent researcher who is blind to intervention condition. Researchers will have prior experience working with young children and will receive BPVS training organised by the ET before the assessment takes place. All researchers will have an enhanced Disclosure and Barring Service (DBS) check and undergo relevant safeguarding and data protection training.

ELIM

Considering the age of targeted children in this trial is 2 years old at baseline assessment and the potential challenges of researcher-administered assessment with children this young, we choose to use one parent-reported language assessment measure-ELIM (at baseline), alongside the researcher-administered language measure (BPVS-III), to fully capture children's vocabulary attainment at baseline. Previous studies have suggested the feasibility of using parent-reported measures such as ELIM in capturing early language skills in young children (Law et al., 2020; Lindorff et al., 2022; Public Health England; 2020).

The ELIM includes a 50-word list developed by Law and colleagues (2020) and has been used previously in England (Public Health England, 2020). It can be used with the parents/carers of children aged between 2 and 2.5 years. It includes words which almost all children at age 2 to 2.5 would be expected to say (for example, mama) and some which only very few children will say (for example, gentle). The full list of words can be seen in the Early language identification measure and intervention: Guidance Handbook published online.

With the authors' permission, an online version of the 50-item vocabulary checklist will be created. The questionnaire will be hosted on a secured digital platform that parents/carers can access via pre-assigned link; they will complete the form in September/October before setting randomisation.

After baseline assessment, a combined outcome of the BPVS-III and ELIM raw scores will be produced to evidence children's receptive and expressive vocabulary attainment before the EYCP intervention. The computation methods of the baseline assessment outcome variable will be discussed in detail in the Study Analysis Plan (SAP).

Primary outcome

Our primary outcome will be children's receptive and expressive vocabulary attainment, as measured by a combination of the researcher-administered BPVS-III and parent/carer-administered MB-CDIs-III at the end of the EYCP intervention.

Similar to the baseline assessment, the BPVS-III (raw scores; 0-168) at post-intervention will be administered one-to-one with children by independent researchers who are blind to intervention condition and meet the same requirements specified above (i.e., have prior experience of working with children; have attended the BPVS training workshops; have an enhanced DBS check; have undergone safeguarding and data protection training).

MB-CDIs-III

The MB-CDIs (Fenson et al., 2007) is a parent report instrument for assessing language and communication skills in children. The complete MB-CDIs measure a range of early communicative and representational skills that are related to language development in typically developing and language-delayed children. Research on the validity of the MB-CDIs and other parent report measures has been generally encouraging and the MB-CDI has a concurrent validity established with the BPVS-III (0.32-0.41) (Baker et al., 2022; EEF, n.d). The MB-CDIs-III (Dale, 2007), a short-form version of the MB-CDIs, was designed to provide more time-efficient and cost-effective instruments to minimise the response burden for parents or the need for highly trained examiners. The validity of parent-reported child vocabulary measures such as the MB-CDI-III, as well as the potential of such measures to be sensitive enough to detect an effect, has been explored in a similar early years study with an RCT design (Sylva et al, 2021); the authors highly recommend the use of a parent-reported child vocabulary measure alongside a researcher or practitioner assessment of child language (Lindorff, Sylva, Ereky-Stevens & Joseph, 2022).The MB-CDI-III 100-item vocabulary checklist section (raw score: 0-100) will be used in this trial to measure children's vocabulary as part of the primary outcome.

With the author's permission, an online version of the 100-item vocabulary checklist will be created for parents/carers to complete in June/July at the end of the EYCP intervention. The questionnaire will be hosted on a secured digital platform and parents/carers will be sent an access link.

A combined outcome of the BPVS-III and MB-CDI-III raw scores will be produced to evidence children's receptive and expressive vocabulary attainment after the EYCP intervention. The computation methods of the primary outcome variable will be discussed in detail in the SAP.

Secondary outcomes

We have two secondary outcomes: 1) children's holistic language and communications skills measured with the Wellcomm Early Years toolkit and 2) practitioner's language-related pedagogical knowledge measured with the OLP tool.

Wellcomm Early Years

The Wellcomm Early Years toolkit is a norm-referenced assessment tool that measures the global/holistic language (Sandwell Primary Care Trust, 2012) of children from 6 months to 6 years old. It was developed by speech and language therapists in the UK and can be used by anyone working with early years children, as it is both relatively short and simple to administer. The toolkit uses a combination of observation, direct assessment, and parent/carer reports to gain a holistic picture of children's language skills. It starts with the 10-question section of the measure corresponding to the child's age (sections cover 6-month age bands). The toolkit uses a traffic light system that bands children by placing them into one of three categories

according to their score (RAG rating): Red: Consider referral to a specialist service for further advice/assessment; Amber: Extra support and intervention required; Green: No intervention currently required. The Wellcomm Early Years Toolkit will be administered by independent researchers after the EYCP intervention, following the same criteria detailed above for the BPVS-III.

The Wellcomm Early Years toolkit raw scores will be used to produce the secondary outcome variable for holistic language and communication skills; the computation methods will be discussed in detail in the SAP.

OLP tool

The OLP tool was developed by Dr Sandra Mathers and findings from the validation study showed that it is a reliable measure of practitioners' procedural knowledge (a secondary outcome in this trial) and that it predicts classroom quality (Mathers, 2020; 2021). It was reported in a recent study (Mathers et al., 2022) that reception class teachers' knowledge as measured by the OLP tool predicted gains in children's understanding of vocabulary and sentence structure between ages 3 and 5. The OLP tool uses videos of real classroom interactions to provide an authentic context for assessing teachers' pedagogical knowledge accurately. It is currently the only video-based tool focused on early language development-related strategies, primarily on adult-child interactions, and it was designed to capture the facets of perceiving, naming and interpreting.

In this trial, the nominated EYCP champions from all settings (both intervention and waitlist-control settings) will be asked to complete the OLP questionnaires online, on a secured platform hosted by the OLP authors/developers, at the end of the intervention in July.

The OLP tool is not commercially available to the public at the moment. We are using the OLP tool with the permission of the author, Dr Sandra Mathers, and she and her team will process the data for the ET in this trial (as the data processor), with practitioners' information being anonymised. The ET team from Durham University will be the data controller in this case, and a data processing and sharing agreement will be signed between the ET and Dr Sandra Mathers.

The OLP tool raw scores will be used to produce the secondary outcome variable for practitioner's language-related pedagogical knowledge; the computation methods will be discussed in detail in the SAP.

Sample size

Given the DT have the maximum capacity to recruit and deliver the EYCP programme to 56 settings per year, the setting sample size is estimated at 112. As such, in this evaluation protocol we calculate the Minimum Detectable Effect Size (MDES) for this two-armed cluster randomised waitlist-control design efficacy trial over 2 years, with 1:1 allocation at the nursery level, based on the following assumptions.

We assume an average of 10 children per nursery at baseline, meaning we will have a total of 1120 children in our sample (112 settings x 10 children per setting), a pre-test/post-test correlation of the primary outcome of 0.60 (i.e., Parentchild+; Tracy et al., 2022; Hanen LLLI; Dimova, et al., 2022) and a nursery-level intra-cluster correlation coefficient (ICC) of 0.15. In an ideal case, without setting-level and child-level attrition, we would have 80% power to detect an effect size of 0.21 of a standard deviation between the control and the intervention groups (see Table 3).

One of our setting eligibility criteria is offering FEE for 2-year-olds, and we assume that 2 or 3 children recruited per setting at baseline are eligible for FEE at age 2. Based on 112 nurseries, with 2 or 3 children with FEE status at age 2 per nursery (total of 224/336 children in an ideal case without attrition), we would have 80% power to show an effect size of 0.32/0.28 of a standard deviation between the control and the intervention FEE subgroups. (see Table 3).

Table 3: Sample size calculations (average 10 children per setting; no attrition)

		Overall	**FEE**
Minimum Detectable Effect Size (MDES)		0.21	0.32/0.28
Pre-test/ post-test correlations	level 1 (pupil)	0.60	0.60
	level 2 (setting)	N/A	N/A
Intra-cluster correlations (ICCs)	level 1 (pupil)	N/A	N/A
	level 2 (setting)	0.15	0.15
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two-sided	Two-sided
Average cluster size		10	2/3
Number of settings	Intervention	56	56
	Waitlist-Control	56	56
	Total	112	112
Number of children	Intervention	560	112/168
	Control	560	112/168
	Total	1120	224/336

One of the setting-level recruitment criteria is having a minimum of 8 eligible 2-year-olds at the start of the school year. However, considering the possible recruitment challenges and in order to meet the setting-level recruitment target (at least 56 settings per school year), the minimum number of eligible children per setting at baseline can be reduced to 6, meaning we would have a total of 672 children in our sample (112 settings x 6 children per setting). Under the same conditions, without setting-level and child-level attritions, we would have 80% power to detect an effect size of 0.23 of a standard deviation between the control and the intervention groups (See Table 4).

Table 4: Sample size calculations (minimum 6 children per setting; no attrition)

		Overall	**FEE**
Minimum Detectable Effect Size (MDES)		0.23	0.32
Pre-test/ post-test correlations	level 1 (pupil)	0.60	0.60
	level 2 (setting)	N/A	N/A
Intra-cluster correlations (ICCs)	level 1 (pupil)	N/A	N/A
	level 2 (setting)	0.15	0.15

Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two-sided	Two-sided
Average cluster size		6	2
Number of settings	Intervention	56	56
	Waitlist-Control	56	56
	Total	112	112
Number of children	Intervention	336	112
	Control	336	112
	Total	672	224

To calculate[5] the MDES, we have used the following formula:

$$D_\sigma = \sqrt{2}\left(z_{1-\alpha/2} + z_{1-\beta}\right)\sqrt{\frac{\left[1+\rho_{icc}\left(m^*-1\right)\right]\left(1-\rho_{cov}^2\right)}{m^*k^*}}.$$

Where

- σ is the MDES in units of standard deviation
- $z_{1-\alpha/2} + z_{1-\beta}$ are the cumulative standard normal values for the significance level (assumed 0.05) and power (assumed 0.8)
- ρ_{icc} is the ICC
- m is the number of children per cluster
- k is the number of clusters per trial arm
- ρ_{cov} is the covariance between baseline and post-intervention assessment results

This calculation assumes a mixed effects model, in which we include the baseline assessment as a covariate. The design effect from the cluster randomised design is included, as well as a factor to account for the reduced variation from including baseline measurements. A detailed MDES calculation process can be found in Appendix F.

Attrition considerations

Attrition has been a concern or risk factor in many EEF early years trials, as stated in the EEF's Lessons Learnt from Early Years Trials report (EEF, 2019), and it is relatively higher in PVI settings due to reasons such as staff turnover. In this efficacy trial, we consider sample attrition at both the setting- and child-level as suggested by other early years trials (i.e., Maths Champions II Effectiveness trial; Robinson-Smith et al., 2023; Hanen LLLI Efficacy Trial; Dimova, et al., 2022).

In this trial, we split 112 settings into two cohorts over 2 years, with 56 settings in each school year (28 in the intervention arm, and 28 in the waitlist-control arm), meaning it will be easier to manage and tackle the attrition issues in comparison to manging all 112 settings in a one-year trial. We aim to implement strategies learnt from the MC Champion II effectiveness trial

[5] Sample size calculations were conducted by a trial statistician (Rachel Oughton) from ET using R (R Core Team, 2022).

(Robinson-Smith et al., 2023) to reduce the setting-level attrition to 5% and the child-level attrition to 10%. Our strategies include:

- Recruiting more SN and maintained nurseries to reduce the overall attrition that might be caused by the high PVI setting attrition.
- Maintaining regular contact with wait-list control settings as well as intervention settings through a key contact-person (i.e., the nominated EYCP champion).
- Gaining opt-in consent from parents/carers at the outset to provide the ET with their child's new setting destination should they leave before the post-test.
- Conducting post-testing at the child's new setting or other location where possible.
- Visiting settings multiple times to complete the child assessment task.
- Supplying a 'thank you' gift to families taking part in the research activities (i.e., a £15 voucher for each round of child assessment tasks completed).

After taking setting- (5%) and child-level (10%) attrition into consideration, the given sample size (112 setting with 10 children per setting at recruitment) would have 80% power to detect an effect size of 0.21 a standard deviation between the control and the intervention groups on the primary outcomes, and an effect size of 0.34 (2 FEE children per setting)/0.30 (3 FEE children per setting) of a standard deviation between the control and the intervention FEE subgroups (see Table 5).

Table 5: Sample size calculations (average 10 children per setting; school-level attrition: 5%; Child-level attrition: 10%)

		Overall	**FEE**
Minimum Detectable Effect Size (MDES)		0.21	0.34/0.30
Pre-test/ post-test correlations	level 1 (pupil)	0.60	0.60
	level 2 (setting)	N/A	N/A
Intra-cluster correlations (ICCs)	level 1 (pupil)	N/A	N/A
	level 2 (setting)	0.15	0.15
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two-sided	Two-sided
Average cluster size		9	2/3
Number of settings	Intervention	53	53
	Waitlist-Control	53	53
	Total	106	106
Number of children	Intervention	477	106/159
	Control	477	106/159
	Total	954	212/318

Randomisation

Nurseries will be randomised to intervention or waitlist-control arm after child recruitment and baseline data collection have been completed. Nurseries allocated to the intervention arm will receive the EYCP programme for free in that school year, those allocated to the wait-list

control arm will continue business as usual during the intervention period and receive the EYCP for free in the following academic year.

For each cohort, an independent trial statistician at Durham University's Mathematical Sciences department will randomise 56 nursery settings using minimisation to ensure balance of sample characters across the trial arms (intervention vs waitlist-control). Minimisation is a valid approach in randomisation especially in small trials (Altman and Bland, 2005) that there will be only minor differences between groups in those variables used in the allocation process.

Minimisation factors[6] that will be considered in randomisation include: 1) nursery type (PVI , maintained and SN settings); 2) child cohort size (<median number of children; >=median number of children); and 3) nursery geographic location (the 8 EYCP mentors are locally sought/based to support 7 settings in a single LA or nearby LAs). The Minim Software is a good resource for running minimisation and will be used in this trial. All nurseries will be informed of their random allocation via a letter emailed to the nursery contact before mid-October.

Statistical analysis

The statistical analysis will follow the most recent EEF Statistical Analysis Guidance (EEF, 2022). All analyses will be conducted on an intention-to-treat (ITT) basis, using two-sided significance at the 5% level. A detailed data analysis strategy will be included in the SAP, to be published in January 2025 (three months after the randomisation of Cohort 2). Below we provide a summary of the planned statistical analysis of primary and secondary outcomes, subgroup analysis, analysis in the presence of non-compliance and additional analyses.

Primary analysis

The primary analysis will investigate any difference in the primary outcome (receptive and expressive vocabulary as evidenced by a combined outcome of BPVS-III with MB-CDI-III) at post-intervention and will take the form of a linear mixed model at the child level with primary outcome score as the dependent variable. Group allocation, cohort[7] and minimisation factors will be included as fixed effects in the model. Adjustments for clustering at the setting level will be made by including the setting as a random effect.

Baseline data will be summarised by trial arm and presented descriptively. Effect sizes based on the difference between the intervention and waitlist-control groups at the post-test will be presented as Hedges' g with 95% confidence intervals. The ICC of the pre-test and post-test for the primary outcome will also be presented.

Subgroup analysis

[6] Nursery type is used as a minimisation factor in randomisation to ensure the type and quality of early years providers are well balanced between the intervention and waitlist-control groups before the intervention; (i.e., staff qualifications and children's disadvantaged status are different between PVI and maintained/SN settings; see Education provision: children under 5 years of age; Gov.UK., 2023); child cohort size is used to ensure the impact of the EYCP is not affected by differences in the staff-child ratio between settings; nursery geographic location is used to align with the EYCP mentoring support mechanism.

[7] While this trial is split into two cohorts of participants, the two cohorts will be analysed as a whole including a covariate for cohort in the model.

The effect of the intervention on the subgroup of children who are 1) eligible for FEE at age 2[8], 2) EAL learners, and 3) least/best performers in the baseline vocabulary assessment[9], will be assessed both via the inclusion of the subgroup status and an interaction term between the subgroup status and allocation in the primary analysis model, and by repeating the primary analysis in each subgroup of children.

Secondary analysis

Secondary analyses will investigate any differences in secondary outcomes - holistic language and communication skills and practitioners' language-related pedagogical knowledge - between groups after the EYCP intervention, using appropriate techniques for the type of data collected (e.g., continuous or categorical data).

In this trial, the raw scores of the Wellcomm Early Years toolkit will be used to represent children's age-related language attainment at post-intervention and this secondary outcome analysis will take the form of a linear mixed model following the same specifications used in the primary analysis. Additional analysis will be considered using the RAG rating as a binary outcome variable (i.e., Green= no intervention needed, Amber/Red: additional support is needed and/or consider referral to a Speech and Language Therapist), taking the form of a mixed-effects logistic regression model, to further investigate if any of the RAG groups are benefitting more/less from the EYCP intervention.

The raw scores of the OLP tool, representing practitioners' total language-related pedagogical knowledge at post-intervention, will also take the form of a linear mixed model following the same specifications used in the primary analysis.

Analysis in the presence of non-compliance

Analysis in the presence of non-compliance will indicate the treatment effects amongst those who participate in the intervention. Compliance in this trial is defined as the extent to which the critical ingredients of the EYCP intervention are delivered to and/or received by the target participants (settings/practitioners/children).

In this trial, compliance will be measured at the nursery level. Each nursery in the intervention arm will be assessed for compliance in terms of the key programme inputs/activities as specified in the EYCP logic model. Table 6 summarises the key EYCP intervention compliance criteria. The exact rating scale of each criterion will be specified in the SAP.

The compliance indicator will first be represented as a continuous variable in the form of a total score for 1) the core components of the EYCP intervention (0-13), 2) the optional components (0-6), and 3) both the core and optional components combined (0-19), as t listed in Table 6. In other words, there will be three continuous indicators to account for nurseries' compliance with the core, optional and all components of the EYCP intervention. Using these indicators, a Complier Average Causal Effect (CACE) analysis will be performed on the primary outcome (vocabulary attainment) to explore the effects of partial compliance. If partial compliance effects exist, two CACE analyses will be conducted defining compliance as a

[8] Considering the upcoming childcare policy reform in England (see Childcare choices; Gov.UK., 2023b) whereby all working families with 2 year olds will receive 15 hours of FEE from April 2024, the Income Deprivation Affecting Children Index (IDACI) rank of the child's postcode (collected from their parent/carer) will be used as a proxy for disadvantage instead of FEE eligibility for Cohort 2 (2024-2025 school year).

[9] The definition of the least and best performers in the baseline assessment of primary outcome will be discussed in the SAP, together with the computation methods for the least/best performer subgroups.

dichotomous variable to estimate bounds for the treatment effects: minimum compliance and optimal compliance. These dichotomous compliance thresholds will be defined empirically and sensitivity analysis comparing the results of different compliance thresholds will be performed (Sagarin et al., 2014).

CACE analyses will use a Two Stage Least Square (2SLS) approach with group allocation as the instrumental variable for the compliance indicator. Results for the first stage (which predicts the compliance indicator using treatment allocation as an instrumental variable alongside all other covariates included in the second stage) will be reported alongside i) the correlation between the instrument and the endogenous variable; and ii) an F test (EEF, 2022).

Table 6: EYCP Compliance Criteria

Description	Core/Optional component	Data collection method
Two nominated professionals (setting manager/EYFS lead and nominated EYCP champion) remain in their EYCP role during the 8-month EYCP intervention	Core	Collected by the DT
The setting receives a minimum of 5 in-setting visits from their mentor	Core	Collected by the DT
The setting undergoes 2 ITERS-3 audits administered by their mentor	Core	Collected by the DT
The setting manager/EYFS lead completes required PD training sessions (1.5 days face to face and 1 pre-recorded session)	Core	Collected by the DT
The setting manager/EYFS lead supports the EYCP champion and programme implementation in their setting	Core	Collected by the DT through a rating scale administered by the DT
The setting manager/EYFS lead accesses EYCP learning materials and online resources	Optional	Collected by the DT
The setting manager/EYFS lead engages in online networking activities	Optional	Collected by the DT
The setting manager/EYFS lead seeks programme support whenever is needed	Optional	Collected by the DT
The EYCP champion completes required PD training led by the DT (1.5 days face to face plus 6 pre-recorded sessions)	Core	Collected by the DT
The EYCP champion attends 9 hub meetings led by their mentor	Core	Collected by the DT (mentors)
The EYCP champion prepares for and responds to in-setting visits delivered by their mentor	Core	Collected by the DT (mentors)
The EYCP champion engages with mentors on how to cascade strategies to other practitioners in their setting	Core	Collected by the DT through a rating scale administered by mentors

Description	Core/Optional component	Data collection method
The EYCP champion accesses programme learning materials and online resources	Core	Collected by the DT (mentors)
Evidence of the EYCP champion using EYCP strategies/techniques in their classroom	Core	Collected by the DT through a rating scale administered by mentors; Collected by the ET through practitioner surveys
The EYCP champion engages in online networking activities	Optional	Collected by the DT
The EYCP champion seeks regular programme support whenever is needed	Optional	Collected by the DT
The EYCP champion cascades EYCP strategies and techniques to wider staff teams	Optional	Collected by the ET through practitioner's surveys
Children's attendance at participating nursery (hours per week; terms during the intervention)	Core	Collected by ET through parents' survey
Children's exposure to EYCP intervention on a daily basis	Core	Collected by the ET through practitioner surveys

Additional analyses and robustness checks

The effect of the intervention on the subgroup of children who are cared for/taught by staff who have low/high self-reported confidence in teaching language, according to the practitioner baseline survey (see IPE design section below), will be assessed both via the inclusion of the subgroup status and an interaction term between the subgroup status and allocation in the primary analysis model, and by repeating the primary analysis in each subgroup of children.

The effect of the intervention on the subgroup of children who are cared for/taught by staff with high qualifications (e.g., <level 4 or above level), according to the practitioner baseline survey, will be assessed both via the inclusion of the subgroup status and an interaction term between the subgroup status and allocation in the primary analysis model, and by repeating the primary analysis in each subgroup of children.

As a robustness check (and quality assurance), we will analyse the primary outcome differences between groups of children that were assessed by different researchers at baseline and post-intervention.

The amount of missing primary baseline and post-test outcome data will be summarised for both arms, and reasons for missing data and patterns of missing data will be explored and provided. The impact of missing data on the primary analysis will be assessed by repeating the analysis on a data set where missing data has been completed using multiple imputations.

In case of differences in programme delivery between the two cohorts (Cohort 1: delivery in the 2023-2024 school year; Cohort 2: delivery in the 2024-2025 school year), we will conduct an alternative analysis of the primary outcome as a robustness check (e.g., analysing the primary outcome for each cohort separately and doing a meta-analysis of the two 'separate cohort trials'.

Implementation and process evaluation design

For more information refer to the EEF's Implementation and Process Evaluation Guidance

The IPE will follow the EEF principles and guidance (EEF, 2022) for undertaking process evaluations. The primary aims will be to monitor EYCP implementation fidelity/adherence and evaluate the components of all stages in the delivery of the intervention, including the training and ongoing PD support, features of optimal implementation, barriers to implementation, adaptations to implementation and perceived progress, whilst considering dosage and reach of programme delivery. Given the EYCP programme will be delivered over two years (three if considering the waitlist control group for Cohort 2), we will use the IPE to explore differences in EYCP delivery between the two cohorts (Cohort 1 in the 2023-2024 school year, and Cohort 2 in the 2024-2025 school year) to inform the interpretation of the impact evaluation findings.

In this trial, we define fidelity as 'the degree to which the (EYCP) intervention is implemented as intended or prescribed' and we explore intervention fidelity from two aspects: fidelity of EYCP PD programme delivery (PD training and support) to settings and fidelity of EYCP programme implementation within intervention settings. In terms of usual practice, we explore the usual practice of all settings (both intervention and control settings) and practitioners in terms of their usual CPD practice (general CPD and language domain-specific CPD) as well as practitioners' usual language teaching and learning-related practice in the setting.

Given this is an efficacy trial, the particular focus of the IPE will be on exploring the programme theory and relationships between different components of the ToC (see Appendix A) and testing the causal and contextual assumptions (see Appendix B and Appendix C). We will gather views from all relevant stakeholders including the EYCP programme developers and DT, nursery managers, nominated EYCP champions, and practitioners via surveys, interviews and/or focus groups and training session observations. As agreed in the set-up meetings with the DT, observations of the setting's learning environment data through the ITERs-3 audit will also be used in the IPE to further understand what 'quality' programme delivery looks like. This IPE will complement the aims of the impact evaluation by *explaining* the observed and perceived effect of the EYCP intervention on children's language and communication development in participating settings.

Research questions

Our main IPE research questions (RQ) and the subgroup questions are listed below.

Fidelity/adherence

RQ1: Is the EYCP PD training and mentoring support delivered as intended to PVI, maintained and SN settings?

1.1. Are nominated staff (manager/EYFS lead, EYCP Champion) attending EYCP training and the hub meetings (EYCP champion) as specified in the programme plan?

1.2. How appropriate are the PD training sessions (e.g., content, coverage, dosage and duration) and the level of support for nominated staff? Are PD training sessions and mentoring support (including hub meetings and setting visits) delivered as manualised?

1.3. What are the barriers and required facilitators for nominated staff to access and engage with the EYCP PD training and mentoring sessions online and/or face-to-face (including hub meetings and setting visits)?

1.4. Is the EYCP PD training and support delivered to settings consistently by EYCP mentors across different hubs?

1.5. To what extent does the EYCP programme and its delivery to settings in Cohort 1 differ from the delivery in Cohort 2?

RQ2: Is the EYCP programme implemented with fidelity in the intervention group in PVI, maintained and SN settings?

2.1. What are the perceived roles of nominated staff (manager/EYFS lead, EYCP Champion) in the programme?

2.2. Do nominated staff to adhere to their roles as requested in the programme?

2.3. To what extent are the EYCP strategies cascaded to the wider team within the setting?

2.4. What are the barriers and facilitators for implementing the programme in PVI, maintained and SN settings?

2.5. What other factors explain variation in the fidelity of implementation in the intervention group?

2.6. To what extent does the EYCP implementation in Cohort 1 settings differ from the implementation in Cohort 2 settings?

Usual practice

RQ3: What is the usual practice in language and communication provision like in all nursery settings?

3.1. What are the key observable features of a setting delivering the EYCP (in terms of interactions with children, etc) in the intervention group?

3.2. To what extent, do the EYCP features differ from the usual language and communication practice in the control group?

RQ4: What is the usual PD practice like in all nursery settings?

4.1. What language-related PD opportunities do staff have in the control group (e.g., DfE's Early Years Education Recovery Programme's Online Early Years Child Development Training strand)?

4.2. What are the perceived impacts of these language-related PD programmes on children's language and communication skills in the control group?

4.3. To what extent, does the EYCP PD programme differ from the usual PD practice in all settings?

4.4. Are nursery staff participating in any other language-related PD programmes (e.g., DfE's Early Years Education Recovery Programme's Online Early Years Child Development Training strand) in the intervention group?

Causal mechanism and contextual assumptions

RQ5: What are the perceived impacts of the EYCP from different stakeholders (including developers, mentors, setting managers, EYCP champions, and practitioners)?

5.1. How far does the EYCP PD training and mentoring support impact the knowledge, skills and confidence of EYCP champions regarding children's language and communication needs?

5.2. What are the observed impacts of the EYCP upon the setting's learning environment as evidenced by the ITERS-3 audit outcome?

5.3. What are the observed impacts of the EYCP upon practitioners' language teaching practice and skills?

5.4. What are the observed impacts of the EYCP on children's language learning as well as other learning areas in the nursery?

5.5. What factors appear to affect the intervention impacts (e.g., EYCP champions' qualification and teaching experience, children's prior language and communication skills, child level exposure of EYCP, FEE status, EAL status)?

5.6 Are settings able to incorporate EYCP strategies into their practice and sustain them in the longer term?

5.7. What are the perceived and observed impacts of the EYCP on disadvantaged children (i.e. with FEE status)? What are the barriers or facilitators that are specific to those disadvantaged children?

5.8. What are the unintended consequences/impacts of the EYCP programme in intervention nurseries?

5.9. To what extent does the perceived impact of in Cohort 1 (2023-2024 school year) differ from the programme impact in Cohort 2 (2024-2025 school year)?

Fidelity, feasibility and quality assurance of impact evaluation

RQ6: To what extent does the EYCP programme impact evaluation process adhere to the plan?

6.1. Does the nursery setting recruitment process adhere to the recruitment strategy?

6.2. Do PVI, maintained and SN settings respond similarly or differently to the recruitment strategy?

6.3. Does the children and family recruitment process adhere to the recruitment plan?

6.4. Do baseline and post-test administrators (independent researchers) effectively and appropriately deliver the assessment tasks?

6.5. To what extent does the setting and child recruitment strategy used for Cohort 1 differ from the strategy used for Cohort 2?

RQ7: What is the feasibility of the impact evaluation plan in this efficacy trial and is it feasible to be scaled up?

7.1. Are parents/carer- administered language measures acceptable for parents/carers and are they feasible to use in a large-scale trial in the future?

7.2. Are setting and child recruitment strategies used in current trial feasible to use in a large-scale trial in the future?

7.3. What challenges/opportunities has the age of the target population (2 to 3 years old) had for delivery of the evaluation?

7.4. What factors may need to be considered in scaling up the intervention to deliver in more nurseries?

Cost evaluation

RQ8: Is the EYCP intervention cost-effective for participating settings?

8.1. To what extent does the EYCP intervention cost for Cohort 1 differ from the cost for Cohort 2?

Good/best practice

RQ9: What are the key features of the EYCP 'good/best practice' settings in current trial?

Research methods

The IPE is designed to be descriptive (cross-sectional and longitudinal) and to be delivered alongside, and complement, the impact evaluation. The IPE will employ different methods to gather data relating to the pre-specified research questions, and the sample size and participants for each element of the IPE study may vary from all participating settings (n=112) to a purposive sample of settings from intervention settings only (n=16, Cohort 1: 6 intervention settings; Cohort 2: 10 intervention settings).

In Cohort 1 (2023-2024 school year), 4 settings will be selected early on (in November after randomisation) to observe the full process of the intervention and to capture different setting characteristics (e.g., setting type, nursery size and location) while making sure that one setting for each paired mentor group (i.e., 8 mentors working in 4 groups) is selected. Another 2 settings will be selected at a later stage, between mid-implementation and end-implementation, to form a 'good/best practice' setting subsample. They will be nominated by the DT based on their observations of settings' engagement with PD training and implementation of the EYCP in settings. It will provide an in-depth exploration of the context in which good/best practice settings are 'successfully' implementing the EYCP programme through a certain type of 'good practice' during the intervention.

In the 2024-2025 school year (Cohort 2), in line with the mentoring arrangement change (i.e. 8 mentors, each to support 7 settings), 8 settings will be selected in November after randomisation to observe the full process of the intervention and to capture different setting characteristics (e.g., setting type, nursery size and location). Like for Cohort 1, 2 settings will be nominated by the DT, between mid-implementation and end-implementation, to form a 'good/best practice' setting sample to explore the context in which good/best practice settings are 'successfully' implementing the EYCP programme in the 2024-2025 school year.

Baseline and post-intervention surveys for practitioners (all settings)

Baseline and post-intervention surveys from relevant staff will collect measures of:

- Confidence and skills in supporting early language development (nominated EYCP champion)
- Confidence and knowledge of the specific EYCP strategies and approaches (nominated EYCP champion)
- Setting usual practice (setting manager/EYFS lead)

- Practitioner background information (setting manager/EYFS lead and nominated EYCP champion)

As the EYCP ToC expects practitioner skill and knowledge in supporting children's language development to improve as part of the programme, it is important to measure this before and after the intervention. To this end, we will use an adapted version of the Practitioner Confidence and Skills in Supporting Children's Language Measure (Lindorff, Sylva, Ereky-Stevens & Joseph, 2022). This measure contains two sections. The first section captures practitioners' confidence in knowledge and skill and asks participants to answer 12 items relating to supporting children's language development. These items use a 5-point Likert scale ranging from 'not at all (1)' to 'very much (5)' in terms of how confident the practitioner feels about each item. The second section contains items on how practitioners interact with individual children. This is measured through 13 items related to practice when interacting with children, looking at how frequently each activity is done using a 5-point Likert scale from 'never (1)' to 'very often (5)' for each item. Additional items will be added to the surveys using the same Likert scale format to capture confidence, knowledge and current practice related specifically to EYCP techniques/approaches (e.g., 'ShREC approach').

The surveys will also contain quantitative and qualitative questions regarding practitioners' usual practice regarding story time and CPD to investigate the context of the delivery of the programme. This will allow the study to look at how EYCP is different to usual practice and the nursery activity that EYCP will be replacing. The baseline practitioner survey will also include items asking about total years of experience in Early Years, highest qualification, type of qualification and role in the nursery.

Administrative records (all intervention settings)

Administrative records (provided by the DT) of practitioner attendance at PD training and support sessions will be used to measure the extent to which settings complied with the PD training and engaged with PD support. Administrative records from EYCP mentors from their mentoring sessions with each setting will capture fidelity to the delivery of the programme's techniques and progress through the 9 hub meetings and 5 nursery visits. It will be used to triangulate findings from other IPE data collection methods.

Observation/audit of nursery learning environment (all intervention settings)

As part of the EYCP programme, ITERS-3 (Infant/Toddler Environment Rating Scale) observations of the setting's environment will be completed by the DT (the mentor) at the beginning and end of the programme in all intervention settings, to provide formative feedback on progress towards providing a language-rich environment. As agreed between the DT and the ET, the ITERS-3 data will be shared with the ET as part of the IPE study, and this will provide insight and clarity into the results of the impact evaluation, and the impact of the EYCP programme on the quality of language provision within nurseries.

Observations of EYCP Training and Mentoring (16 intervention settings)

Two half-day observations of the face-to-face EYCP training for all practitioners will be conducted by two researchers from the ET, attending the full sessions using a semi-structured observation schedule. The observation schedule will be developed by the ET in consultation with the DT. This will investigate whether the EYCP PD trainings are delivered as manualised and to what extent PD training sessions are adhering to a pre-specified plan for the session.

Observations of 8 hub meetings delivered by the mentors will be conducted cross-sectionally by two researchers from the ET using a semi-structured observation schedule between

January and June in each intervention year. This will investigate whether the EYCP mentoring support scheme is delivered as manualised. Observations of hub meetings will investigate the extent to which hub meetings are adhering to a pre-specified plan for the session and whether the mentoring support is delivered with consistency between mentors.

For quality assurance, we will use inter-rater reliability (IRR) methods to assess the consistency and agreement between the observers. To assess IRR, there are different statistical methods such as percentage agreement, Cohen's kappa, Fleiss' kappa, and ICC. These methods help to quantify the degree of agreement between the observers and provide a measure of the consistency of observations. We will calculate both percent agreement and Cohen's kappa[10].

Interviews and/or focus groups (16 intervention settings)

Semi-structured interviews and/or focus groups[11] with practitioners (setting managers/EYFS lead, EYCP champions, and a third practitioner when applicable[12]) from intervention settings will be undertaken cross-sectionally between mid-implementation and end-implementation of the EYCP (either face-to-face or online via a video conferencing software). These will investigate how the EYCP training has been delivered, how the setting has implemented the EYCP programme through the year, the fidelity of delivery, adherence to the implementation plan, any difficulties with implementing the programme, as well as gathering practitioner views of the acceptability and impact of the programme.

Semi-structured focus groups with the DT (including the ELIS team and the mentor team), either face-to-face or online, will capture their views of the EYCP programme delivery across the project, any challenges they faced, any changes they made to the programme during delivery and how actual delivery compares to their expectations. Focus groups with the DT will be undertaken longitudinally at two-time points: one in January/February at the mid-implementation stage, and one in July/August when the intervention is completed.

Semi-structured focus groups with the ET will be conducted to capture the impact evaluation adherence to the evaluation plan and the quality assurance process[13] after the intervention in August/September.

All interviews and focus groups will be audio-recorded for transcription and quality assurance purposes. Interviews and focus group practice will be conducted by two researchers separately, and a third researcher will observe two live interview/focus group sessions per researcher and monitor the interviews/focus group sessions throughout the process for quality assurance purposes. The third researcher will sample 20% of interview/focus group transcripts and check against the video recording to identify mistakes, such as missed and misheard words, in order to provide an independent interpretation of the information content.

Analysis

[10] If there is likely to be much guessing among the observers, it may make sense to use Cohen's kappa; but if observers are well-trained and little guessing is likely to exist, we may safely rely on percent agreement to determine IRR (McHugh, 2012).

[11] A focus group will be used when all practitioners from one setting are only available for limited interview timeslots.

[12] If the EYCP champion is claiming to be cascading EYCP strategies to the wider team, a third practitioner will be invited to interview.

[13] An independent researcher from the School of Education, Durham University who is not directly working for this trial will conduct the focus group with evaluators XQ, VM, NS, RO and the other two research officers who are supporting the impact evaluation and quality assurance process.

Interviews and/or focus group data

All data collected in interviews and focus groups will be analysed using thematic and content analysis and will be initially coded deductively based on the project's research questions and then inductively looking for any unexpected emerging themes. We will follow the stages outlined by Braun and Clarke (2006): detailed familiarisation; generating initial codes; searching for themes; reviewing themes; defining and naming themes and data reporting. The original transcripts will be coded deductively to develop an initial coding framework based upon the specified IPE dimensions and research questions, and the development and refinement of the coding framework will be iterative throughout analysis.

Survey data

Descriptive statistics will be presented for the quantitative survey questions showing the exploration of data by randomised group, setting type, practitioner role and practitioner experience as well as by time point (where relevant). Continuous measures will be reported as a mean and standard deviation, while categorical data will be reported as a count and percentage.

Data from the 'early language development confidence and knowledge' and the 'confidence and knowledge of the specific EYCP techniques/approaches' measures will be analysed using a linear mixed model, adjusting for the nursery-level minimisation factors (nursery type, nursery child cohort size and nursery location), cohort and highest qualification of the respondent as a fixed effect, and nursery as a random effect.

Qualitative survey responses will be coded deductively using thematic analysis in line with the research questions.

Observation data

Observation data from EYCP training and mentoring sessions will be analysed deductively to look at the level of consistency in the delivery of training sessions between mentors and the engagement of participants.

ITERS-3 audit/observation data

Descriptive statistics will be presented for all intervention settings to explore the differences between two-time points: the beginning of the intervention and the end of the intervention. As ITERS-3 will be administered by the DT (mentors), the independency of the data and findings should be taken into careful consideration.

Results from all IPE data collection methods (observations, interviews/focus groups, surveys, administrative data) will be triangulated and synthesised from the themes and presented as answers to the IPE research questions under the main IPE dimensions. Table 7 summarises the data collection methods used to address the IPE research questions.

Table 7**:** IPE design, methods of data collection and analysis overview

Research methods	Data collection methods	Participants data sources (type, number)	Data analysis methods	IPE research question (RQ)	Implementation / logic model relevance
Longitudinal design (between settings analysis)	Semi-structured focus group	DT team (from ELRS) (n=3) and mentors (n=8); At two time points: mid-, and end-points of implementation	Combination of inductive and deductive analysis, with analyses grouped thematically according to RQs	RQ1: 1.1-1.4 RQ2: 2.2 RQ5: 5.1-5.2; 5.8-5.9 RQ6: 6.1-6.3;6.5 RQ7: 7.2-7.3. RQ9	IPE/logic model; Fidelity/ adherence; Context
Cross-sectional design (between settings analysis)	Semi-structured Interview and/or focus group	Setting managers/ EYFS leads (n=16)	Combination of inductive and deductive analysis	RQ1: 1.1-1.3 RQ2: 2.1-2.2; 2.4-2.5 RQ3: 3.1-3.2 RQ4: 4.1-4.4 RQ5: 5.3; 5.5-5.8 RQ6: 6.1-6.3; RQ7: 7.3-7.4	IPE/logic model; Fidelity/ adherence; Inputs/outputs; Process; Outcomes; Context; Cost evaluation
Cross-sectional design (between settings analysis)	Semi-structured Interview and/or focus group	EYCP champions (n=16)	Combination of inductive and deductive analysis	RQ1: 1.1-1.3 RQ2: 2.1-2.5 RQ3: 3.1-3.2 RQ4: 4.1-4.4 RQ5: 5.1; 5.3-5.8 RQ6: 6.1-6.3; RQ7: 7.3-7.4	IPE/logic model; Fidelity/ adherence; Inputs/outputs; Process; Outcomes
Cross-sectional design (between settings analysis)	Semi-structured Interview and/or focus group	Other relevant practitioners	Combination of inductive and deductive analysis	RQ2: 2.1-2.5 RQ3: 3.1-3.2 RQ4: 4.3-4.4 RQ5: 5.4-5.8	IPE/logic model; Fidelity; Inputs/outputs; Outcomes; Context

Research methods	Data collection methods	Participants data sources (type, number)	Data analysis methods	IPE research question (RQ)	Implementation / logic model relevance
Cross-sectional design (between settings analysis)	EYCP training session observations	DT with EYCP champions and/or setting managers/EYFS leads from intervention settings 2 half-day observations of PD training 8 Hub meeting observations (1-hour)	Descriptive/thematic analysis; Content analysis	RQ 1: 1.1-1.5	IPE/Logic model; Fidelity/adherence; Input/output.
Longitudinal design (between settings analysis)	Administrative records of training attendance and engagement	All intervention settings (n=56)	Descriptive analysis; Regression analysis	RQ1: 1.1; 1.5 RQ7: 7.3-7.4	IPE/Logic model; Compliance; Context
Longitudinal design (between settings analysis)	Nursery learning environment observation (ITERS-3)	All intervention settings (n=56)	Descriptive analysis	RQ5: 5.2; 5.9	IPE/Logic model; Fidelity/adherence; Input/output; Context; Outcomes
Longitudinal design (between settings analysis)	Baseline and endpoint usual practice surveys	All control and intervention nurseries (n=112)	Descriptive analysis; Content analysis; Thematic analysis	RQ3: 3.1-3.2 RQ4: 4.1-4.4	IPE/Logic model; Usual practice; Context; Compliance
Longitudinal design (between settings analysis)	Baseline and endpoint practitioners' confidence surveys	All control and intervention nurseries (n=112)	Descriptive analysis; Liner regression analysis	RQ5 5.1;5.3; 5.5; 5.8	IPE/Logic model; Outcomes; Context
Longitudinal design	Cost evaluation workshops	DT team (n=3)	Cost table for ingredients of intervention	RQ8: 8.1	Cost evaluation
Longitudinal design	Baseline and endpoint nursery cost evaluation survey	All intervention settings (n=56) Setting managers/EYFS leads; EYCP champions	Descriptive analysis	RQ8: 8.1	Cost evaluation
Longitudinal design	Semi-structured focus group	Evaluators from ET	Thematic analysis	RQ6: 6.1-6.5 RQ7: 7.1-7.4	Feasibility and quality assurance

Research methods	Data collection methods	Participants data sources (type, number)	Data analysis methods	IPE research question (RQ)	Implementation / logic model relevance
Cross-sectional design (between settings analysis)	Observation of child assessment sessions	ET assessors for child assessment (n=>23)	Descriptive analysis; Content analysis; IRR analysis: the percentage agreement and Cohen's kappa	RQ6: 6.4 RQ7: 7.3	Feasibility and quality assurance
Good/best practice case study (within settings analysis)	Utilising all data collection methods listed above	DT nominated settings between mid- and end-implementation (n=4)	Thematic analysis; Grounded theory approach	RQ9	IPE/Logic model; Context; Good practice

Cost evaluation

The cost evaluation will follow the most recent guidance from the EEF (2023). All cost analyses will be conducted from the perspective of education services (i.e., nurseries) (principle 1). The costs associated with programme implementation will rely on the 'ingredients method' principle (Levin et al., 2018; EEF, 2023) and all resources necessary to implement the EYCP programme will be accounted for. The EYCP programme costs will be divided into 'pre-requisites, start-up costs and recurring costs' (principle 7). The primary sources of cost are expected to be staff costs for training and the cost of resources, and these will be incurred at two key time points, specifically at training and during implementation. The cost evaluation will estimate average marginal costs per pupil per year over three years (principle 8).

Sources of data

The costs in this trial will be estimated using two different data sources:

1. Data provided by the EYCP DT on costs related to delivering PD training (including face-to-face and virtual sessions), hub meetings, in-school support visits (including the ITERS-3 audits at two-time points) and any other costs associated with providing support to settings as part of the EYCP programme.

2. Data on the costs of setting participation such as staff cover for attending EYCP training sessions online or face-to-face, accommodating mentors' school visits, as well as the cost of resources for EYCP programme implementation within settings.

Data collection methods

Cost data will be collected from relevant staff members by the ET at different time points throughout the trial. The collection of this cost data will be integrated into the IPE data collection methods.

- Cost workshops. Three semi-structured workshops will be conducted with the DT to establish the cost of delivering the EYCP intervention: one early on to establish the ingredients model to be used after the intervention plan is finalised and encourage the collection of cost data, one after Cohort 1 implementation and before the Cohort 2

implementation to capture ongoing costs and to plan for Cohort 2 cost data collection, and one towards the end of the Cohort 2 implementation to capture costs for the whole trial over two years.

- EYCP intervention cost-specific surveys. Two rounds of EYCP cost-specific surveys will be conducted with all intervention settings for each Cohort. The first one will capture the amount of time (staff working hours) that are spent completing the relevant training components of the programme, any start-up, prerequisites costs (e.g., computer or internet connectivity), recurring costs associated with training (e.g., storybooks), as well as any other costs associated with the implementation of the EYCP (e.g., to cover staff absence). The second one will capture the amount of time (staff working hours) that are spent implementing the programme, continuing to deliver the programme and any recurring implementation costs (e.g., materials, printouts, resources) and unexpected or hidden costs.
- IPE interview/focus groups. During the IPE interview, setting managers/EYFS leads will also be asked specific questions relating to the cost of the EYCP intervention. This may be used to detect any unexpected costs across various settings and to provide an in-depth exploration of the context in which possible cost variations are occurring between settings.

Ethics and registration

Ethics

Ethics approval for this trial was initially obtained from the Durham University Ethics Committee on 3rd March 2023. Programme changes regarding sample size and lengths of intervention were obtained ethical approval on 29th August 2023.

Registration

This trial will be registered on the Open Science Framework (OSF) site in September/October 2023.

Data protection

The lawful basis we are relying on is 'public task', which according to General Data Protection Regulations (GDPR) Article 6 (1) (e) means 'the processing is necessary for an activity being carried out as part of the University's public task, which is defined as teaching, learning and research'. This project is carrying out research and will not be processing special category personal data.

A Data Protection Impact Assessment (DPIA) has been conducted for this project by the principle investigators, describing why the collection of personal data is necessary, the data that will be collected and the ways that risks will be minimised. A data management plan is also in place to set out the procedures for the team to follow to ensure data confidentiality throughout the project. A three-way DSA: controller to controller will be put in place between the DT at ELRS, the ET at Durham University and all nurseries to ensure that all teams handle data correctly and that data can be shared securely and appropriately. A DSA: controller to processor will be put in place between ET at Durham University and Dr Sandra Mathers at Oxford University as data processor for the OLP tool, to ensure that data can be handled, processed and shared securely and appropriately.

All information relating to relevant children, staff and nurseries in this trial will be treated in the strictest confidence and processed in accordance with GDPR. Durham University is the data controller and the data processor (except for the OLP data and the administrative data collected by DT) in this evaluation trial. Anonymous data may be kept indefinitely by the ET and potentially shared with other research teams. The lawful basis for processing data would be 'Public Task'.

We will provide a privacy notice for settings, practitioners and parents/carers which describes in detail how we will collect, store, process and share setting, practitioner, child and parents/carers' data as part of the EYCP evaluation study. At the end of the evaluation study, data will be submitted to the EEF's data archive. At this point, the EEF and its archive manager (FFT Education) are responsible for controlling and processing the evaluation study data.

Personnel

Delivery Team

Fliss James: is Director of East London Research School based at Sheringham Nursery School and Children's Centre. Fliss is joint programme delivery lead and has responsibility for the design, delivery and management of the project.

Melissa Prendergast: is Deputy Headteacher of Sheringham Nursery School and Children's Centre. She strategically leads East London Research School and A Brighter Start: East London's Stronger Practice Hub. Melissa is joint programme delivery lead and has responsibility for the design, delivery and management of the project.

Evaluation Team

Dr Xiaofei Qi: is an Assistant Professor at the School of Education at Durham University. Xiaofei is the lead principal investigator and has overall responsibility for the evaluation design and protocol development, delivery, reporting and overseeing the project management.

Vic Menzies: is an Assistant Professor at the School of Education at Durham University. She is the joint principal investigator and has oversight of the trial design, delivery, and reporting. She is responsible for data management and the cost evaluation.

Dr Nadia Siddiqui: is a Professor at the School of Education at Durham University. She is a co-investigator and is responsible for quality assurance of the evaluation, and will also contribute to a subgroup analysis of pupils from disadvantaged backgrounds.

Dr Nashwa Ismail: is a Lecturer in Education at the School of Education at Durham University. She is a co-investigator (until September 2023) and is responsible for conducting elements of the IPE study.

Dr Rachel Oughton: is an Associate Professor in Statistics at the Department of Mathematical Sciences at Durham University. She is a co-investigator and is responsible for the development of the SAP, and will also oversee the statistical analysis for this trial.

Dr Dandan Chen: is a Research Associate at the School of Education, at Durham University. She is the research support officer managing the EYCP project and with a specific responsibility for the Implementation and Process Evaluation of the project.

Rowan Van-Muysen: is a Research Assistant at the School of Education, at Durham University. He is the project Coordinator liaising with various stakeholders and with a specific focus on the impact evaluation and data management of the evaluation data.

Risks

Table 8: Risks

Risk	Mitigation
Insufficient settings recruited	• The ET will work closely with the DT to deliver recruitment events. • Longer period of recruitment. • Randomisation in batches. • Recruit PVI, maintained and SN settings. • Widening target regions and areas.
Insufficient children recruited	• Use EoI sheets to gather eligible children information at setting recruitment stage. • Consider widening the age group of targeted children (e.g., 2 to 4 year olds) • Provide parents/carers with quality information about the trial/intervention. • Enhance recruitment strategies (e.g., use of pre-registration lists). • Parents/carers assured of confidentiality of data.
Baseline data collection	• Minimise setting level burden in baseline data collection. • Incentives for settings in supporting baseline assessment. • Use E-consent forms and digital forms for data collection when participants have easy access and feel comfortable doing so. • Incentives for parents/carers in completion of baseline assessment.
High attrition (settings)	• DT and ET to develop strong relationship with settings through regular contact. • Reiterate the value of waitlist-control nurseries in initial communications. • Aim to over recruit to allow for unavoidable attrition. • EEF prepares a letter to the setting managers/ EYFS leads to encourage they remain in the trial.
High attrition (children)	• Maintain regular contact with settings to follow up on children's attendance. • Identify children who are leaving the setting at an early stage. • Develop a follow-up plan with parent/carers and consider offering extra incentives to keep in contact with families for post-intervention assessment.
Post-test data collection	• Offering an incentive to the setting following completion of the post-test. • Use family incentives (e.g., 'thank you' for completion of post-tests). • Accommodate settings' needs while arranging child assessment activities. • Include consent to follow and test children at new destination or alternative location. • Multiple visits to settings for child assessment.
Missing data	• Setting characteristics, current practice surveys, participating child details and baseline assessments will be required as a condition to be randomised. • Outcome measures selected based on minimal burden for nurseries.
Cross-over	• Children may move from an intervention setting to a control setting or vice versa. Children's data will be analysed as per the original assignment (ITT) and cross-over considered within a CACE analysis.
Nursery staff turnover	• The trained EYCP champion may leave the setting leading the setting to disengage. The setting manager/EYFS lead or another practitioner within the setting could mitigate this risk by taking over the lead role.

Timeline

Table 9: Timeline

Dates	Activity	Staff responsible/ leading
Nov 22-Jan 23	Trial set up	ET, DT and EEF
Feb-Sept 23	Protocol and ethics	ET
EYCP Trial in Cohort 1: 2023-2024 School Year		
Feb-Sept 23	Recruitment of settings	DT (with ET support)
Sept-Oct 23	Recruitment of settings and parent/carers	DT (with ET support)
July-Sept 23	Recruit and train assessors for child language assessment at baseline	ET
Sept-Oct 23	Pre-intervention (baseline) surveys for nursery practitioners	ET
Sept-Oct 23	Pre-test for children	ET
Oct 23	Randomisation	ET
Nov 23-Jun 24	Delivery of the EYCP	DT
Nov/Dec 23	IPE developer interview with DT	ET and DT
Nov-Dec 23	IPE Observations I	ET
Jan-April 24	IPE Observations II; practitioners interview/focus group	ET
May-July 24	IPE Observations III; practitioners interview/focus group	ET
May-Jun 24	Recruit and train assessors for child language assessment at endline	
Jun-Jul 24	Post-test for children; Post-intervention surveys for nursery practitioners	ET
Aug/Sept 24	IPE developer interview with DT	ET and DT
EYCP in Cohort 2: 2024-2025 School Year		
Oct 23-Sep 24	Recruitment of settings	DT (with ET support)
Sept-Oct 24	Recruitment of settings and parent/carers	DT (with ET support)
July-Sept 24	Recruit and train assessors for child language assessment at baseline	ET
Sept-Oct 24	Pre-intervention (baseline) surveys for nursery practitioners	ET
Sept-Oct 24	Pre-test for children	ET
Oct 24	Randomisation	ET
Nov 24-Jun 25	Delivery of the EYCP	DT
Nov/Dec 24	IPE developer interview with DT	ET and DT
Nov-Dec 24	IPE Observations I	ET
Jan-April 25	IPE Observations II; practitioners interview/focus group	ET

Dates	Activity	Staff responsible/ leading
May-July 25	IPE Observations III; practitioners interview/focus group	ET
May-Jun 25	Recruit and train assessors for child language assessment at endline	
Jun-Jul 25	Post-test for children; Post-intervention surveys for nursery practitioners	ET
Aug/Sept 25	IPE developer interview with DT	ET and DT
Sept 25-Jul 26	Delivery of intervention to Cohort 2 waitlist settings	DT
Analysis and reporting		
Sept 25-Jan 26	Data analysis, report writing	ET
Jan 26	Submit evaluation report	ET
May 26	Publication of evaluation report	EEF
May 26	Data Archiving	ET

References

Borm, G. F., J. Fransen, and W. A. Lemmens. 2007.. A simple sample size formula for analysis of covariance in randomized clinical trials. *Journal of clinical epidemiology 60*(12), 1234–1238.

Baker, S., Harding, S., Holme, C., Lewis, R., Seifert, M. and Wren, Y. 2022. Review of Early Language Screening Suitable for Children in Wales from Birth to 5 Years. Cardiff: Welsh Government, GSR report number 55/2022. Available at: https://gov.wales/early-language-screening-review.

Braun, V. and Clarke, V. 2006. 'Using thematic analysis in psychology', Qualitative Research in Psychology, 3(2), pp. 77–101. DOI: 10.1191/1478088706qp063oa.

Burchinal , M.R, Pace A, Alper R, Hirsh-Pasek K, Golinkoff RM. 2016. *Early language outshines other predictors of academic and social trajectories in elementary school*. Presented at Assoc. Child. Fam. Conf. (ACF), Washington, DC, July 11–13.

Campbell, M., J. Grimshaw, N. Steen, and C. P. P. in Europe Group (EU BIOMED II Concerted Action) .2000. Sample size calculations for cluster randomised trials. *Journal of health services research & policy 5*(1), 12–16.

Dale, P.S. 2007. The MacArthur-Bates Communicative Development Inventory – III. Baltimore, MD: Brookes Publishing.

Dimova, Sashka, Scott, Molly, Scandone, Berenice, Sciarra, Alessandra, Rezaian, Mansor ., Bury, Jonah. 2022. Hanen Learning Language and Loving It (Hanen LLLI) Evaluation protocol. London: Education Endowment Foundation.

Department for Education (DfE). 2023. Statutory framework for the early years foundation stage. Avaliable at: https://www.gov.uk/government/publications/early-years-foundation-stage-framework—2. London: Department for Education.

DfE. 2023. Early years stronger practice hubs Guidance. Available at: https://www.gov.uk/government/publications/early-years-stronger-practice-hubs/early-years-stronger-practice-hubs. London: Department for Education.

Dunn, L.M., Dunn, D.M. and National Foundation for Educational Research (NfER), 2009. The British Picture Vocabulary Scale – Third edition. London: GL Assessment.

Dunn, L.M., and Dunn, L.M., 1981. PPVT-R (Peabody picture vocabulary test – revised) Manual. Circle Pines, MN: American Guidance Service.

Eadie, P., Bavin, E., Bretherton, L.., Cook, F., Gold, L …Reilly, S.. 2021. Predictors in Infancy for Language and Academic Outcomes at 11 Years, *Pediatrics*, 147 (2). e20201712. 10.1542/peds.2020-1712.

Education Endowment Foundation. n.d. Early Years Toolkit: Communication and language approaches. London: EEF Retrieved from: https://educationendowmentfoundation.org.uk/education-evidence/early-years-toolkit/communication-and-language-approaches.

Education Endowment Foundation. n.d. Early Years measures database. Review of measures for children aged 0 – 6 in language, literacy, numeracy and social and emotional skills. London: EEF Retrieved from: https://educationendowmentfoundation.org.uk/projects-and-

evaluation/evaluation/eef-outcome-measures-and-databases/early-years-measures-database-2/early-years-measures-database

Education Endowment Foundation [EEF]. 2022. Statistical Analysis Guidance for EEF Evaluations. London: EEF.

Education Endowment Foundation [EEF]. 2023.Cost evaluation guidance for EEF Evaluations. London: EEF.

Education Endowment Foundation [EEF].2022. Implementation and process evaluation guidance for EEF evaluations. London: EEF.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, J. S. 2007. MacArthur-bates communicative development inventories. Baltimore, MD: Brookes Publishing.

Gelman, A. and J. Hill. 2006.. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gov.UK. 2023a. Data set 2 Children registered and providers by provider type. Education provision: children under 5 years of age, Reporting year 2023 –Explore education statistics. Available at https://explore-education-statistics.service.gov.uk/data-tables/fast-track/4891da30-91d7-44b5-75ef-08db786e904b. (last accessed: 31/07/2023). London: HM Government.

Gov.Uk., 2023b. Upcoming changes to childcare support. Available at: https://www.childcarechoices.gov.uk/upcoming-changes-to-childcare-support/#:~:text=The%20upcoming%20expansion&text=This%20means%20that%3A,to%203%2Dyear%2Dolds. (last accessed: 31/07/2023). London: HM Government.

Harms, T., Cryer, D., Clifford, R. M., & Yazejian, N. 2017. Infant/Toddler Environment Rating Scale, third edition. New York, NY: Teachers College Press.

Haux, Tina., Butt, Sarah., Rezaian, Mansor., Garwood, Eliza., Woodbridge, Hannah., Bhatti, Sehaj., Richard, Woods Rogan, and Paul, Gillian. 2022. The early years workforce: recruitment, retention, and business planning. Department for Education. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1072062/SCEYP_thematic_report-_April_2022.pdf. (Last accessed 24/05/2023).

Hoff, E. 2013. Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Dev. Psychol.* 49, 4–14. DOI: 10.1037/a0027238.

James., F. 2022. The ShREC approach – Four evidence informed strategies to promote high quality interactions with young children. EEF blog. London: EEF. Available at https://educationendowmentfoundation.org.uk/news/the-shrec-approach-four-evidence-informed-strategies-to-promote-high-quality-interactions-with-young-children (Last accessed 30/05/2023).

Law J, Charlton, J, Asmussen, K .2017. Language as a child wellbeing indicator. London: The Early Intervention Foundation.

Law, J. Charlton, J. Dockrell, J. Gascoigne, M. McKean, C. Theakston, A. 2017. Early Language Development: Needs, provision, and intervention for preschool children from socioeconomically disadvantaged backgrounds. EEF Report. London: EEF.

Law, J., Charlton, J., McKean, C., Watson, R., Roulstone, S., Holme, C., Gilroy, V., Wilson, P., Rush, R. (2020). Identifying and Supporting Children's Early Language Needs. Newcastle, UK: Newcastle University.

Lehrl, S., Evangelou, M., and Sammons, P. 2020. The home learning environment and its role in shaping children's educational development. Sch. Eff. Sch. Improv. 31, 1–6. DOI: 10.1080/09243453.2020.1693487.

Levin, H. M. et al., 2018. Economic Evaluation in Education. Cost-Effectiveness and BenefitCost Analysis (Third Edition). London: Sage Publishing.

Lindorff, Ariel., Sylva, Kathy., Ereky-Stevens, Katharina and Joseph, Allen .2022.. Coaching Early Conversation, Interaction and Language (CECIL) Impact evaluation. Sutton Trust. Retrieved from https://www.suttontrust.com/wp-content/uploads/2022/03/CECIL-Evaluation-Oxford-14.02.2022.pdf.

Mathers, S. 2020. Observing Language Pedagogy (OLP): Developing and piloting a contexualised video-based measure of early childhood teachers' pedagogical language knowledge. Doctoral thesis (Ed. D), UCL (University College London).

Mathers, S. J. 2021. Using video to assess preschool teachers' pedagogical knowledge: explicit and higher-order knowledge predicts quality. Early Childhood Research Quarterly, 55, 64–78. DOI: 10.1016/j.ecresq.2020.10.010.

Mathers, S., Malmberg, L.E., Lindorff, A.M. & Gardiner, J. 2022. Early educators' knowledge of early language pedagogy: How can it be measured and does it matter for child language outcomes? Oxford: University of Oxford.

Melhuish, E.C., Phan, M.B., Sylva, K., Sammons, P., Siraj-Blatchford, I. and Taggart, B. 2008. Effects of the Home Learning Environment and Preschool Center Experience upon Literacy and Numeracy Development in Early Primary School. Journal of Social Issues, 64: 95-114. DOI: 10.1111/j.1540-4560.2008.00550.

Mol, S. E., and Neuman, S. B. 2014. Sharing information books with kindergartners: the role of parents' extra-textual talk and socioeconomic status. *Early Childhood Res. Q.* 29, 399–410. DOI: 10.1016/j.ecresq.2014.04.001.

McHugh, M. L. 2012. Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276–282.

Pace, Amy., Luo, Rufan/., Hirsh-Pasek, Kathy., and Golinkoff, Roberta Michnick. 2017. Identifying Pathways Between Socioeconomic Status and Language Development. Annual Review of Linguistics. 3:1, 285-308.

Public Health England. 2020a. Best start in speech, language and communication: Guidance to support local commissioners and service lead. Public Health England. vailable at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/931310/BSSLC_Guidance.pdf.

Public Health England. 2020b. Early language identification measure and intervention: Guidance handbook. London: Public Health England, Department of Health & Social Care, Department for Education.

R Core Team. 2022.. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Robinson-Smith, Lyn, Whiteside, Katie Torgerson, Carole Fairhurst, Caroline Podmore, Danielle, Davill, Tom, Bell, Kerry Qi, Xiaofei Leggett, Heather, Elliott, Louise Hewitt, Catherine Baird, Kalpita, Menzies, Victoria, Torgerson, David & Ainsworth, Hannah. 2023. Independent evaluation of Maths Champions in nursery to develop children's early numeracy: A two-armed cluster randomised controlled trial. Education Endowment Foundation.

Sagarin, B. J., West, S. G., Ratnikov, A., Homan, W. K., Ritchie, T. D., & Hansen, E. J. 2014. Treatment noncompliance in randomized experiments: Statistical approaches and design issues. Psychological Methods,19, 317–333. DOI: 10.1037/met0000013.

Sénéchal, M., Ouellette, G., & Rodney, D. 2006. The misunderstood giant: On the predictive role of vocabulary to reading. In S. B. Neuman & D. Dickinson (Eds.) *Handbook of Early Literacy* (Vol.2; pp.173-182). New York: Guilford Press.

Snow, C.E., Burns, M.S., & Griffin, P. 1998. *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Schwab, JF, Lew-Williams, C. 2016. Language learning, socioeconomic status, and child-directed speech. Wiley Interdiscip Rev Cogn Sci. 7(4):264-275. DOI:10.1002/wcs.1393.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. 2004. *The effective provision of pre-school education (EPPE) project technical paper 12: The final report-effective pre-school education*. IOE: London.

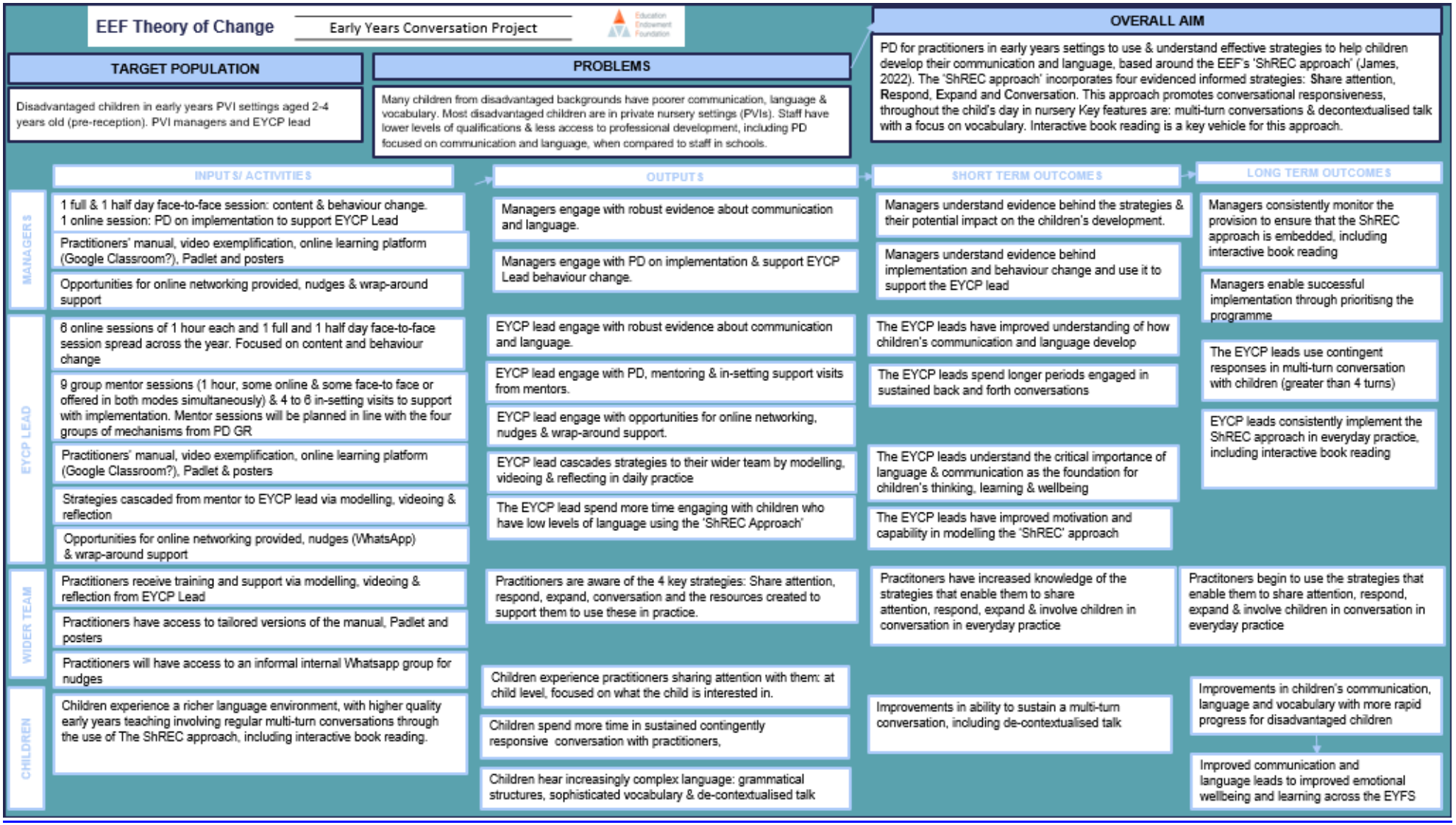
Tracey, Louise, Torgerson, Carole, Welch, Charlie, Fairhurst, Caroline, Gridley, Nicole & Qi, Xiaofei. 2019. Efficacy of the ParentChild+ programme: trial evaluation protocol. Education Endowment Foundation.

Tracey, Louise., Torgerson, Carole., Dysart, Erin., Fairhurst, Caroline., Gridley, Nicole. & Welch, Charlie. 2022. Efficacy Trial of the ParentChild+ programme: Evaluation Report. Education Endowment Foundation.

WellComm. 2010. A Speech and Language Toolkit for Screening and Intervention in the Early Years. GL Assessment. Retrieved from: https://www.glassessment.co.uk/assessments/products/wellcomm.

Wooles, N., Swann, J., & Hoskison, E. 2018. Speech and language delay in children: a case to learn from. *The British journal of general practice : the journal of the Royal College of General Practitioners*, *68*(666), 47–48. DOI: 10.3399/bjgp17X694373.

Appendix A: EYCP Theory of Change



Appendix B: Contextual assumptions log

#	Contextual Assumption	Assumption Strength	Assumption Risk
1	Managers and EYCP champions can access the online training and attend all of it; they can attend the 1.5 days of face-to-face training which will require cover	Green	Red
2	EYCP champions can access the online 1-hour group mentor sessions and attend all 9 hub meetings	Green	Amber/Green
3	Managers and EYCP champions have gaps in their knowledge or hold misconceptions about the most effective ways to support children's language development	Green	Amber/Red
4	EYCP champions are motivated to engage in the training	Amber/Green	Amber/Red
5	EYCP champions have time to implement the strategies in their practice, without other priorities taking over	Amber/Green	Amber/Red
6	EYCP champions have the support of their managers to implement the strategies	Amber/Green	Amber/Red
7	Practitioners are motivated to engage in the training from their room lead	Amber/Green	Amber/Red
8	EYCP champions have time to train their team in using the strategies	Amber/Green	Red
9	EYCP champions have the necessary skills to monitor consistent implementation, model strategies and offer feedback to all their team-members	Amber/Green	Red
10	Settings will accommodate the minimum 5 visits to support with implementation	Green	Amber/Red
11	Children spend more time in sustained conversation with practitioners	Amber/Green	Red
12	Children hear a wider and richer range of vocabulary	Amber/Green	Red
13	Children from socioeconomically disadvantaged background (i.e., FEE) benefit more from the programme		

Assumption strength

Green – This assumption will hold in the vast majority of circumstances where the programme is delivered

Green/Amber – This assumption will hold in most of the circumstances where the programme is delivered

Red/Amber – This assumption will often not hold in the circumstances where the programme is delivered

Red – There is a good chance of this assumption not holding / do not know whether this assumption will hold or not

Assumption risk

Green – The programme could continue to be delivered with very minor impact

Green/Amber – The programme could continue to be delivered, but the impact would be substantial

Red/Amber – The programme could continue to be delivered, but without fidelity to original design

Red – The programme could not be delivered

Appendix C: Assumption/causal mechanism

#	Assumption / Causal Mechanism	Where in the ToC does the assumption apply?	Evidence	Evidence Strength
1	Practitioners engage in training, mentoring and in-setting support visits	Output	Robust evidence: engaging in PD which includes a range of mechanisms (building knowledge; motivating; developing teaching techniques; embedding practice) EEF PD GR; FEEL study	Green
2	Practitioners engage with online networking, nudges, and wrap-around support	Output	Robust evidence: engaging in PD which includes a range of mechanisms (developing teaching techniques - social support; embedding practice - providing prompts and cues, prompting action planning) EEF PD GR; FEEL study	Green
3	Practitioners engage with robust evidence about communication and language	Output	Robust evidence: motivating (presenting information from a credible source) - Law et al, Ang and Harmey, and other strong sources of evidence about children's needs and effective strategies	Green
4	Children experience practitioners sharing attention with them, to develop conversation.	Output	Law et al, Ang and Harmey and other strong sources of evidence about children's needs and effective strategies - however evidence also suggests that achieving precise and sustained changes in practice is difficult to achieve	Amber/Green
5	Children spend more time in sustained conversation with practitioners, including decontextualised talk	Output	Law et al, Ang and Harmey, Romeo et al and other strong sources of evidence about the importance of conversation, including decontextualised talk; however evidence also suggests that achieving precise and sustained changes in practice is difficult to achieve	Amber/Green
6	Children hear a wider and richer range of vocabulary	Output	Law et al, Ang and Harmey and other strong sources of evidence about the importance of vocabulary; we are not aware of a robust study testing this approach (focus on hearing richer range of vocabulary) for children of this age	Amber/Red
7	Practitioners have improved understanding of how children's communication and language develop	Short-term outcome	Robust evidence: engaging in PD which includes a range of mechanisms (building knowledge; motivating; developing teaching techniques; embedding practice) EEF PD GR; FEEL study; the evidence about how children's communication and language develops is extensive and secure	Green
8	Practitioners' skills in interacting with children improve	Short-term outcome	Robust evidence about effective strategies to use, but there are no studies which test this particular combination of strategies and their impact	Amber/Green

9	Practitioners' understand the critical importance of language and communication as the foundation for children's thinking and learning	Short-term outcome	Law et al, Ang and Harmey and other strong sources of evidence about the importance of communication and language in early childhood	Green
10	Improvements in ability to sustain a multi-turn conversation, including decontextualised talk	Short-term outcome	Evidence is growing in strength (e.g., Romeo et al) but not yet secure and we are not aware of robust studies which consider this in the context of group settings	Amber/Red
11	Improvements in vocabulary	Short-term outcome	Robust evidence - importance of vocabulary; importance of the 'language-rich' environment with an extended, not restricted code; some research demonstrates a causal link between practitioners using a wider and richer vocabulary, and children improving their vocabulary at a faster rate	Amber/Red
12	Practitioners consistently implement the ShREC strategies in everyday practice	Long-term outcome	The ShREC strategies are based on robust evidence; however, the impact of using this specific set of strategies has not been tested; evidence also suggests that consistent implementation is difficult to achieve	Amber/Red
13	Improvements in children's communication and language	Long-term outcome	Robust evidence suggests that using a range of evidence-informed strategies effectively will improve children's language development; however, many children with below-expected language in the early years will naturally 'catch up' without the need of any different practices; many settings access PD around communication e.g., from ICAN; establishing this causal link is likely to be difficult	Amber/Green
14	Improved communication and language leads to improved emotional wellbeing and learning across the EYFS	Long-term outcome	Robust evidence suggests that communication is the foundation of thinking, play and learning. Children with more extensive vocabularies learn to read faster; however, establishing a causal link will be challenging - lack of robust measures for these items; impact of other aspects of pedagogy	Amber/Red

Evidence Strength - How strong is your evidence base?

Green – The evidence base is very strong. There are peer-reviewed academic studies, meta-analyses or independent experimental evaluations directly linked to the assumption.

Green/Amber – The evidence base is strong. There are academic studies or independent evaluations linked to the assumption.

Red/Amber – The evidence base is developing. There are academic studies, internal evaluations or recorded observational evidence that are adjacent to the assumption.

Red – The evidence base is limited. There is some anecdotal evidence to substantiate the assumption. You might not be aware of any evidence linked to the assumption.

Appendix D: Changes since the previous evaluation

Feature		Pilot to efficacy stage
Intervention	**Intervention content**	The previous intervention was also focussed on conversational responsiveness techniques, but there were more included than the 'ShREC' focus of the EYCP intervention as delivered in the efficacy trial. Additional techniques included: • OWL (Observe, Wait, Listen): rather than talking, an adult should observe what a child is doing, wait for them to talk, and listen to what the child says; • The four finger rule: every question should be followed by three comments to extend conversations with a child; • 'Wait, watch, wonder': use a child's spontaneous activity or free play to start a conversation.

	Delivery model	There are a number of changes to the delivery model:
		• All practitioners in each setting were trained and expected to deliver the techniques, whereas in the efficacy trial there is a single EYCP champion who is trained alongside the setting manager.
		• The training was structured as follows:
		  ○ Initial launch event for managers
		  ○ Bespoke training at each setting
		  ○ Monthly coaching sessions
		  ○ Termly mentoring sessions
		• Training was carried out by independent consultants rather than the core ELRS team, and the mentoring was carried out by mentors from the project team rather than the independent coaches.
		• In the pilot the audit tool was optional.

	Intervention duration	The previous intervention ran from Sept 2018 to June 2019. As such, the duration of delivery was very similar to the EYCP programme delivery: it has not been shortened but just shifted along a bit.
Evaluation	**Eligibility criteria**	Recruitment was limited to the London Borough of Newham, and only PVI settings were eligible to participate.
	Level of randomisation	Not applicable to pilots.
	Outcomes and baseline	Not applicable to pilots.
	Control condition	Not applicable to pilots.

Note: This document was written by Melissa Prendergast and Fliss James from the DT, and with revisions from the ET.

Appendix E: EEF-SPH recruitment strategy template

This is a basic template to help grantees provide some key information about their recruitment strategy. Feel free to add sections and extra information. Please ensure this is sent to and discussed with the EEF prior to recruitment starting.

- **Target number of settings** to be recruited (as agreed in set up meetings)
- **Selected recruitment areas**
- **How many eligible settings are there** in these areas (what is your target pool?)
- **Are there any other EEF trials recruiting** similar settings during the same time frame (ask your Programme Manager if you are not sure)
- **Key activities and methods.** Please detail here key recruitment activities, for example recruitment events, any system leaders you will contact, methods for approaching schools and promoting the project
- **What is the timeline** for recruitment? Will you set targets throughout the time period?
- **How many staff do the team have dedicated to the recruitment phase** of the project and in what capacity? How will you ensure everyone is trained up to support recruitment?
- **Who in the setting will be targeted**?
- **What is the process from a setting expressing an interest to signing the MOU?** How will you monitor this internally? Will you set deadlines for schools to return the MOU?
- **How will you ensure settings have a full understanding** of what is involved?

Appendix F: Sample size calculations

(Rachel Oughton October 2022)

This summarises the method used to calculate samples size / MDES.

Equations used

The numbers in the sample size table 3-5 were calculated using information in Campbell et al. (2000), Chapter 20 of Gelman and Hill (2006) and Borm et al. (2007). These work in terms of total number of people recruited for a two-arm trial, rather than number recruited per arm. This led to problems when rearranging the formula to solve for MDES. So, in what follows, *n* is the number of participants per trial arm, and *k* is the number of clusters per trial arm. The particular equation was:

$$n = \left[\frac{2\sigma^2\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{D^2}\right]\left[1 + \rho_{icc}\left(m - 1\right)\right]$$

Where

- σ is the standard deviation in the population outcome
- α is the type I significance level
- β is the type II significance level (so $1 - \beta$ is the power)
- D is the minimum detectable effect size (MDES)
- ρ_{icc} is the ICC
- m is the average cluster size

If the effect size is thought of as a multiple of the standard deviation,

$$D = \sigma D_\sigma$$

then this simplifies to

$$n = \left[\frac{2\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{D_\sigma^2}\right]\left[1 + \rho_{icc}\left(m - 1\right)\right] \quad (1)$$

Accounting for correlation with a covariate

Since we can assume a moderate correlation between the baseline and follow-up measurements, we can use the results of Borm et al. (2007) and include the baseline measurement in a linear regression model for the follow-up measurement to achieve the same power with a sample size reduced by a factor of $1 - \rho^2$, compared to that in Equation (1).

Therefore our sample size calculation becomes:

$$n = \left[\frac{2\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{D_\sigma^2}\right]\left[1 + \rho_{icc}\left(m - 1\right)\right]\left(1 - \rho_{cov}^2\right) \quad (2)$$

where ρ_{cov} is the correlation between baseline and follow-up measurements.

Rearranging for MDES

Because we know the number of settings is fixed due to DT capacity, it is best to rearrange the formula to solve for MDES, to see what effect size we can achieve under given conditions. Rearranging Equation (2) gives:

$$\begin{aligned} D_\sigma &= \sqrt{\left[\frac{2\left(z_{1-\alpha/2}+z_{1-\beta}\right)^2}{n}\right]\left[1+\rho_{icc}\left(m-1\right)\right]\left(1-\rho_{cov}^2\right)} \\ &= \sqrt{2}\left(z_{1-\alpha/2}+z_{1-\beta}\right)\sqrt{\frac{\left[1+\rho_{icc}\left(m-1\right)\right]\left(1-\rho_{cov}^2\right)}{n}}. \end{aligned} \quad (3)$$

Cluster numbers and attrition

Rather than the whole number of participants *n*, we are more concerned with the number of clusters, $k = n/m$. Therefore, Equation (3) becomes

$$D_\sigma = \sqrt{2}\left(z_{1-\alpha/2}+z_{1-\beta}\right)\sqrt{\frac{\left[1+\rho_{icc}\left(m-1\right)\right]\left(1-\rho_{cov}^2\right)}{mk}} \quad (4)$$

Finally, we need to take attrition into account. If we assume attrition rates of γ_k at the setting/cluster level and γ_m at the participant or child level, then our effective (or final) number of children per setting becomes:

$$m^* = m(1 - \gamma_m)$$

and the effective (or final) number of settings **per arm** becomes:

$$k^* = k\,(1 - \gamma_k).$$

So, if we recruit *k* settings with *m* children in each at the start, we can expect our MDES D_σ to be approximately:

$$D_\sigma = \sqrt{2}\left(z_{1-\alpha/2}+z_{1-\beta}\right)\sqrt{\frac{\left[1+\rho_{icc}\left(m^*-1\right)\right]\left(1-\rho_{cov}^2\right)}{m^*k^*}} \quad (5)$$

Calculated values

Values of MDES (D_σ) have been calculated in R (R Core Team, 2022) using Equation (5), and are shown in Table 1. The fixed values are:

$\gamma_k = 0.15$

$\gamma_m = 0.2$ $\alpha = 0.05$ $\beta = 0.2$ (therefore power = 0.8)