Education Endowment Foundation

Evaluation summary

Project title	**Making it REAL evaluation**
Developer *(Institution)*	National Children's Bureau (NCB)
Evaluator *(Institution)*	National Centre for Social Research
Principal investigator(s)	Dr Svetlana Speight
Protocol author(s)	Svetlana Speight, Maha Basharat, Helena Takala, Konstantinos Papaioannou, Alina Fletcher, Ekaterina Stoilova
Trial design	Two-arm cluster randomised control trial with random allocation at setting level
Trial type	Efficacy
Pupil age range and Key stage	Early years, aged 3-4
Number of schools *(at design stage)*	150
Number of pupils *(at design stage)*	900
Primary outcome measure and source	Receptive language (The British Picture Vocabulary Scale (BPVS3))
Secondary outcome measure and source	Receptive and expressive language (The Renfrew Action Picture Test (RAPT))

Protocol version history

Version	Date	Reason for revision
1.2 [*latest*]		
1.1		
1.0 [*original*]		N/A

Contents

Tables

Figures

Study rationale and background

Background

Policy context

The importance of foundational literacy skills for children's development has been reflected in a concerted policy focus on early literacy in England over the past decade. The statutory framework for Early Years Foundation Stage (EYFS) (DfE, 2021b) recognises children's literacy as one of seven key areas of learning and development in the early years. The framework highlighted the importance of developing children's life-long love of reading, consisting of language comprehension and word reading. It also notes that the development of children's spoken language underpins all seven areas of development.

Home learning environment[1] has been a focus of policy interest over the past decade. This is because a key factor mediating children's literacy development is the number of words and quality of language engagement to which a young child is exposed (Law et al., 2018). Given that children spend considerable time in the parental home in their early years, there has been substantial interest among policy-makers and researchers in what happens in the home regarding communication, interaction and the role of parents in improving their children's literacy outcomes.

In 2018, the DfE published a policy paper outlining a behaviour change approach aimed at improving the home learning environment (DfE, 2018). The paper recognised that there are barriers to supporting children's early language and literacy development in the home, particularly in disadvantaged areas. These include the capability of parents, the availability of opportunities, and motivational barriers. It suggested sharing messages and resources with families in disadvantaged areas and using positive behaviour role models to emphasise to parents what they can easily implement in the home, as ways of overcoming these barriers.

More recently, policy and research attention has focused on the negative impacts of the COVID-19 pandemic on early literacy. An EEF study reported that reception class children's EYFS attainment levels were lower than expected in the 2021/2022 cohort compared with what was achieved in the 2018/2019 cohort (Tracey et al., 2022). There also continues to be a gap in the attainment of socially disadvantaged children and their more advantaged peers: in 2022, less than 50% of children eligible for free school meals were found to leave reception class achieving a 'good level of development' (on the EYFS measure in 2021) (DfE, 2022a). In 2023, the government pledged to support the recovery from the COVID-19 pandemic in relation to children's early literacy and language attainment (DfE, 2023). This included dedicating funds for the early years sector to upskill practitioners to enable them to better support children's early literacy and language development.

[1] The 'home learning environment' includes the physical characteristics of the home (e.g., the availability of resources such as books, a quiet place to study, and access to IT equipment) as well as the activities parents/caregivers engage in with children at home, which support a child's learning (e.g., reading stories, reciting nursery rhymes, drawing, learning numbers and shapes).

Existing evidence

Existing research shows that parental involvement in children's learning is associated with a range of positive impacts on children's cognitive and socio-emotional outcomes (Asmussen et al., 2016; Melhuish and Gardiner, 2018), as well as with impacts on academic achievement (Lehrl et al., 2020; McNeal, 2014; Melhuish and Gardiner, 2020). The quality of the home learning environment has also been shown to be as important as socioeconomic factors (Melhuish and Gardiner, 2018). Several studies have found that positive parental support for home learning has a greater impact on children's wellbeing and achievement than any other factor (Axford et al., 2015; Siraj-Blatchford and Siraj-Blatchford, 2009).

The original Sheffield REAL Project was developed in the 1990s by Professor Cathy Nutbrown and Professor Peter Hannon from the University of Sheffield, who later worked with the National Children's Bureau (NCB) to develop the Making it REAL programme. REAL stands for Raising Early Achievement in Literacy. Making it REAL aims to help practitioners (teachers and nursery nurses) to build parents' knowledge and confidence in creating a home learning environment that supports and encourages their children's early literacy development.

The REAL approach has been previously evaluated. Two evaluations that included randomised controlled trials showed the approach to be effective (Hirst et al., 2010; Hannon et al., 2020). A third evaluation showed modest but consistent increases in child-reported family literacy activities (Nutbrown and Hannon, 2003). In addition, there has been a feasibility trial of the Making It REAL programme that was carried out as part of the Sutton Trust's Engagement Fund in 2015–2017 (Sylva et al., 2018).

In 2018, EEF commissioned NatCen to conduct the first large-scale evaluation of the REAL programme. The evaluation was originally designed as a two-armed cluster (setting-level) randomised control efficacy trial involving 120 nurseries. However, the impact evaluation was cancelled in March 2021 due to challenges of continuing with the trial in the light of the COVID-19 pandemic and the subsequent implications for programme implementation and data collection. Instead, NatCen carried out a standalone implementation and process evaluation (IPE) with all settings to assess implementation and perceived outcomes on practitioners, parents, and children.

This evaluation is the first efficacy trial of the Making it REAL programme. The evaluation design is informed by lessons learnt from the previous REAL programme evaluation (Speight et al, 2022).

Integrated evaluation design

This evaluation comprises an impact evaluation and an IPE. It will run over two years, with two distinct cohorts of children (Cohorts 1 and 2). The aim of the impact evaluation is to establish the impact of Making it REAL on children's receptive and expressive language skills, compared to children who receive usual early learning and care. The aim of the IPE is to understand how the intervention was delivered in practice and to provide context for interpreting findings from the impact evaluation. In addition, the IPE will explore differences between Cohorts 1 and 2, to see whether the programme delivery evolved and whether this has implications for the interpretation of findings from the impact evaluation.

Intervention

Name

Making it REAL

Who (recipients and provider)

Teachers and early years practitioners will deliver the intervention to children aged 3- to 4-years-old. Private, voluntary and independent (PVI) early years settings, maintained nursery schools (MNS) as well as nursery classes attached to primary schools will be eligible to participate in the trial. NCB will be responsible for recruiting settings, delivering the Making it REAL training to practitioners and supporting settings and practitioners with delivery.

NCB will recruit 150 settings over two years (75 settings in Cohort 1 and 75 settings in Cohort 2) to take part in the evaluation. Six children in each setting will take part in baseline and endline testing. In settings that are randomised to the treatment group, families will receive the Making it REAL intervention. In settings that are randomised to the control group, families will receive the setting's usual early learning and care.

Settings will nominate two practitioners per treatment setting to attend the training run by NCB and to deliver the intervention with families. At least one out of two practitioners must be qualified to level 3 or higher to be eligible to deliver Making it REAL.

What (materials and procedures)

Making it REAL is based on the Opportunities, Recognition, Interaction and Modelling (ORIM) framework, which highlights four approaches parents can take to create a home learning environment that supports children's early literacy development:

1. Creating **O**pportunities for children's literacy development, for example by making children books, CDs and writing material available and accessible in the home environment;

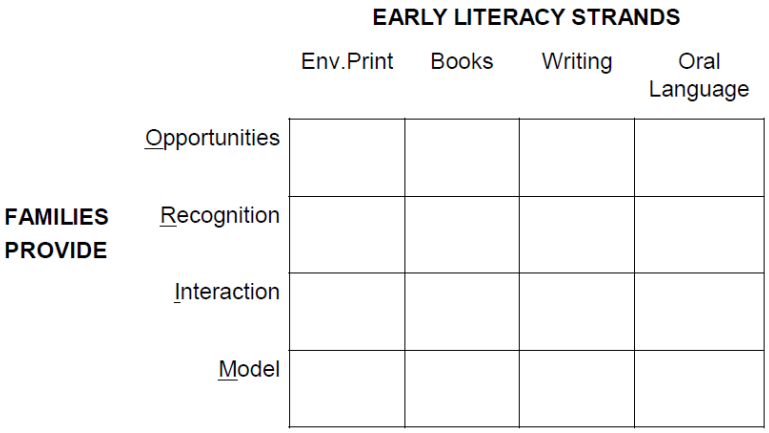
2. **R**ecognising and encouraging children's literacy milestones;

3. **I**nteracting with children positively and supporting real-life literacy tasks and;

4. Acting as **M**odels of literacy users, so children see parents use literacy in everyday life.

The ORIM Framework distinguishes four inter-related strands of emerging literacy: environmental print, books, early writing, and key aspects of oral language (storytelling, phonological awareness, and 'language for literacy') as shown in Figure 1 (Nutbrown et al., 2005; Hannon, 1995).

Figure 1: The ORIM Framework



Early years practitioners delivering the Making it REAL programme will be required to attend a two-day training course delivered by NCB. The training will cover: 1) approaches to developing children's literacy in current academic literature, 2) how to understand and interpret the ORIM framework and use it with families, and 3) how to work with families in the home and overcoming common barriers to support this. Practitioners will receive a training pack containing the ORIM framework grid, suggestions for activities and resources for home visits and literacy events, research sheets, DVDs containing examples of home visits and literacy events, and other planning materials.

The trained practitioner will carry out four **home visits** to each family. Home visits are guided by the ORIM framework grid, which the practitioner completes and that they might share with families. During the home visits, practitioners will model activities and encourage parents (or carers/guardians) to mirror the activities with their child outside the home visit. In the following visit, practitioners will encourage parents to reflect on their experiences implementing the activities covered in the previous home visit. At the end of the visit, practitioners leaves parents with resources and instructions for activities to carry out with the child.

Practitioners also arrange four **literacy events**, open to all parents of children aged 3- to 4-years-old in the nursery. The literacy events are a further opportunity for practitioners to introduce and engage parents with the four strands of literacy and offer parents a range of activities to try in the home. In addition, literacy events provide a further opportunity for modelling – both from practitioner to parent, practitioner to child, but also parent to parent as parent or carer observes other families interacting over a literacy activity.

Lastly, practitioners are expected to attend four **network meetings,** which will be organised and facilitated by NCB. The aim of the network events is to create a space for trained practitioners to network and share their experiences of delivering Making it REAL. They provide an opportunity to celebrate successes, but also to jointly consider barriers and solutions, to be reminded of the training content, and share tips and ideas for visits and events.

Where

The trial will take place in around five local authorities in Yorkshire and the Humber. These local authorities will support NCB with the recruitment of settings by utilising their contacts with local settings, providing information on the programme, and generating interest in taking part.

Home visits will take place in the child's home. Literacy events will be held in the nursery setting, or at an alternative local venue such as a library, a park, or a museum. NCB will set the location and timing of in-person practitioner training and network events.

How

NCB associates will deliver practitioner training in person. Network meetings will be delivered either in person or online using a video conferencing software (e.g., Zoom).

Practitioners will deliver all homes visits and literacy events in person. They will use the ORIM framework grid to plan activities to share with families. The child and one parent (or carer or guardian) must be present at each visit. Practitioners can bring new resources (e.g., books, toys, stationary) to help model activities to families, or use objects borrowed from the nursery setting or school. They can also use objects found within the home.

When and how much (Dosage)

NCB will deliver training for practitioners at the end of the autumn term. Practitioners will carry out home visits and literacy events over spring and summer terms. There will be two cohorts taking part in the evaluation. Cohort 1 settings will take part during the 2023/24 school year, and Cohort 2 settings will take part during the 2024/25 school year.

Practitioner training will last two full days consecutively. Practitioners will also attend four network meetings during spring and summer school terms, each lasting around one hour.

Practitioners will deliver four home visits to each child during the spring and summer terms. Each visit should last from 40 minutes to an hour. Practitioners will be asked to space out home visits (i.e., two home visits per each school term) to give parents sufficient time to successfully implement activities that practitioners have modelled to them.

During the spring and summer term, practitioners will also organise four literacy events for parents, which will each last around one hour. Practitioners can decide the sequencing of home visits and literacy events throughout the school year – i.e., they can deliver a literacy event before a home visit, or vice versa.

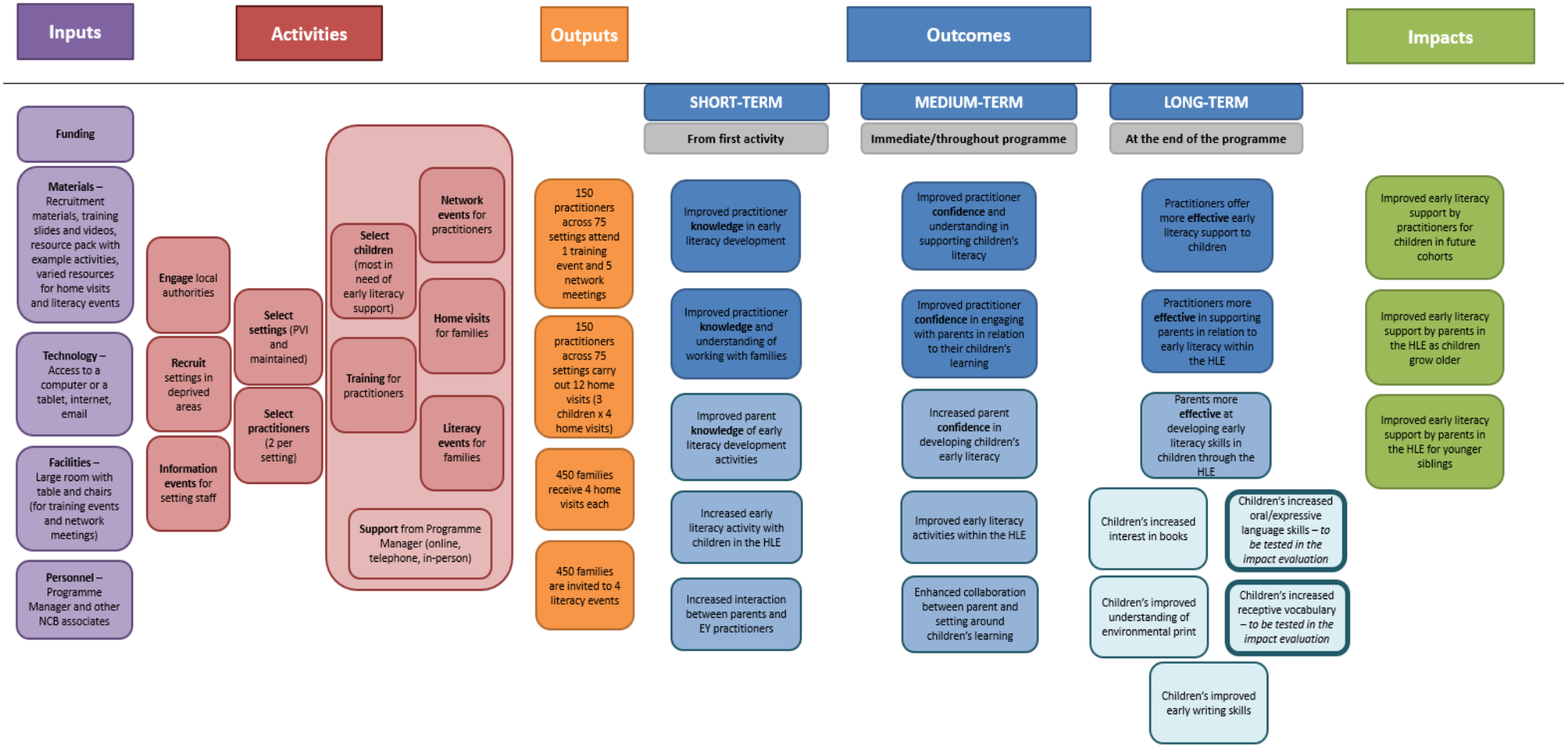
Tailoring (Adaptation)

Practitioners are encouraged to use the ORIM framework in designing home visits and literacy events, but there is room for flexibility in which resources they use and which activities they suggest to the families. Practitioners can be reflective and reactive in designing the structure and content of home visits and literacy events. During training, they are encouraged to tailor the programme activities to the needs of the individual children and their families.

Logic model

Figure 2 displays the intervention logic model.

Figure 2: Making it REAL Logic Model



Impact evaluation design

Research questions

The impact evaluation of Making it REAL will address the following research questions:

- RQ1: What is the impact of the Making it REAL programme on 3- to 4-year-old children's receptive language as measured by the British Picture Vocabulary Scale (BPVS3), compared to children who receive usual early learning and care? (Primary outcome)
- RQ2: What is the impact of the Making it REAL programme on 3- to 4-year-old children's receptive and expressive language as measured by the Renfrew Action Picture Test (RAPT), compared to children who receive usual early learning and care? (Secondary outcome)
- RQ3: Does the impact of the Making it REAL programme on 3- to 4-year-old children's receptive language, as measured by the BPVS3, differ by Early Years Pupil premium status? (Subgroup analysis)

Design

Table 1 provides a summary of the trial design.

Table 1: Trial design

Trial design, including number of arms		Two-armed cluster randomised trial
Unit of randomisation		Setting
Stratification variables (if applicable)		Geographic area, setting type (school-based nursery class/MNS vs group-based PVI)
Primary outcome	**Variable**	Receptive language
	Measure (instrument, scale, source)	British Picture Vocabulary Scale (BPVS3) score; age standardised, 70-140, GL Assessment
Secondary outcome(s)	**Variable(s)**	Receptive and expressive language
	Measure(s) (instrument, scale, source)	Renfrew Action Picture test (RAPT) score; 5th Edition Information and Grammar 2020 raw score, 0-41 Information and 0-39 Grammar, Routledge
Baseline for primary outcome	**Variable**	Receptive language
	Measure (instrument, scale, source)	British Picture Vocabulary Scale (BPVS3) score; age standardised, 70-140, GL Assessment
Baseline for secondary outcome	**Variable**	Receptive language
	Measure	British Picture Vocabulary Scale (BPVS3) score; age standardised, 70-140, GL Assessment

	(instrument, scale, source)	

The impact evaluation of Making it REAL is designed as a two-arm cluster randomised control efficacy trial, with settings as the unit of randomisation and children as the unit of analysis.

The trial will employ setting-level randomisation to prevent the risk of potential spillover effects associated with individual-level randomisation within settings. A total of 150 settings will be recruited to participate in the trial across two years (Cohort 1 and Cohort 2). Settings will be stratified by local authority (geographic area) and setting type (school-based nursery class/MNS vs group-based PVI setting) and randomised into a treatment or a control arm within each stratum (see Randomisation

All participating settings will have a 50:50 chance of being assigned to the treatment or control group within each area-setting-type stratum. Settings assigned to the treatment group will receive training and deliver Making it REAL home visits and literacy events, as described in the Intervention section above. Those assigned to the control group will continue with their usual early learning and care. Random assignment ensures that, on average across trials, there are no differences at baseline between intervention and control. This makes it possible to estimate the average causal effect of Making it REAL on outcomes.

The trial will run as a two-year split-cohort trial. A total of 150 settings will be recruited across the two years: 75 in year 1 (Cohort 1) and 75 in year 2 (Cohort 2). Cohort 1 settings will be recruited and randomised into the treatment or control arm (resulting in 37 or 38 settings in each arm). Control settings from this cohort will be eligible to sign up to take part in Cohort 2 and be randomised when we run the study again with a new set of settings. However, it is possible that they could be randomised to control again.

All participating settings will be asked to identify children eligible to receive the Making it REAL programme (see Participants). Financial incentives will be offered to both treatment and control settings after baseline testing (£150) and after endline testing (£150) in both years to help promote involvement in the trial and limit attrition.

It is possible that programme delivery differs between Cohort 1 and Cohort 2 due to external factors and growing delivery team expertise. The implementation and process evaluation (see Implementation and process evaluation (IPE) design) will aim to explore whether the programme delivery evolved and whether this has implications for the interpretation of findings from the impact evaluation.

The selected outcome measures are age-appropriate[2], fit well with two of the long-term outcomes in the logic model of the programme and were selected in collaboration with the delivery team. The primary outcome of interest is receptive language as measured by BPVS3. The secondary outcome of interest is receptive and expressive language as measured by

[2] We will use the standardised BPVS3 score for the primary outcome, constructed using conversion tables available in the BPVS3 manual. The standardised scores reflect how a child's score compares to other children of a similar age in months, based on a large nationally representative standardisation sample. The RAPT test provides a raw RAPT score because this test has been standardised only for children who speak English as a first language.

RAPT. Both outcomes together provide a comprehensive measure of children's language development and are described in more detail in the Outcome measures

section below.

Participants

Study participants

The Making it REAL programme will be delivered by early years practitioners (teachers and early years educators/practitioners) to children aged 3 to 4 in nursery settings.

Inclusion and exclusion criteria

Setting eligibility

Both school-based settings (nursery classes in state primary schools and maintained nursery schools) and group-based providers in the PVI sector (such as day nurseries and pre-schools) are eligible to participate in the evaluation of Making it REAL. The inclusion of PVI settings introduces some important considerations for the trial (e.g., power), which are discussed later in this section and in the Sample size calculations

section. Settings randomised to the control group in Cohort 1 will be eligible to sign up for Cohort 2. There are no eligibility criteria around the number of hours a child attends the setting or the number of hours the setting is open. However, all settings must have a minimum of 10 children in the eligible age group (aged 3 in September 2023 for Cohort 1) in order to take part in the study. All settings will be recruited from around five local authorities in the Yorkshire and the Humber region.

Practitioner eligibility

Settings will allocate two named early years practitioners to the project to work with NCB. This could be a teacher, a teaching assistant, an early years educator or another staff member working directly with children. At least one of the two staff members selected must be qualified to Level 3 or higher.

Child eligibility

Six children per setting will take part in the evaluation, three children per practitioner. Practitioners will select eight children aged 3 to take part in the programme in September 2023 for Cohort 1 and September 2024 for Cohort 2. They will select children they assess to be most in need of early literacy support (NCB will provide further guidance on child eligibility; however, the programme is aimed at those children most likely to not meet a good level of development at Early Years Foundation Stage Profile). Only six of the eight children will take part in baseline assessments. The remaining two are considered as reserve children. This will help mitigate risks of attrition due to child absence or late parent opt-outs and ensure that six children can be tested on the day of baseline assessments. Should all eight children be present on the day of the assessments, NatCen will randomly select six children to take part in assessments and two children to not take part.

Parents

To be eligible, parents must not have opted out of consent. All parents of eligible children will be contacted. NatCen will provide an information leaflet for the setting and a parent information sheet along with the project's privacy notice so that all nurseries and parents are fully aware of the requirements of the evaluation.

Recruitment strategy procedures, roles, and responsibilities

The total number of planned treatment units across Cohort 1 and Cohort 2 is 75. Recruitment is scheduled for April-July 2023 for Cohort 1 and January-July 2024 for Cohort 2.

We expect NCB to identify and recruit eligible schools and EY settings and help them select two eligible practitioners to take part, with NatCen advising on eligibility criteria and communicating requirements for research participation. NCB will in the first instance collaborate with local authority early years teams to gain their support to access settings. Local authorities will facilitate communications with settings, advise on strategy for local recruitment, and where possible provide encouragement to settings to take part.

All participating settings will be required to complete a Memorandum of Understanding (MoU). Settings agreeing to participate in the trial will be asked to allocate two eligible practitioners to the project to work with NCB. Practitioners will then be asked to identify and select children eligible to take part. NatCen will provide a trial information sheet and privacy notice for parents or carers of eligible children to inform them about the trial.

As a potential benefit of the programme spanning two years, the recruitment strategy can be improved for Cohort 2, e.g., in the case of concerns about attrition in Cohort 1.

Outcome measures

Primary outcome

The primary outcome of interest is English language ability, measured using the British Picture Vocabulary Scale third edition (BPVS3). BPVS3 is a test of receptive vocabulary for Standard English suitable for children 3-16 years old and for those with learning difficulties. A child's receptive vocabulary is tested by asking children to identify pictures that illustrate a given word's meaning. Each item comprises an array of four colour illustrations. The child's task is to select the picture that best illustrates the meaning of a stimulus word spoken by the test's administrator. For example, the administrator would say the word "ball" and the child would have to identify the picture illustrating a ball among an array of pictures illustrating a rose (picture 1), a pumpkin (picture 2), a ball (picture 3) and a parrot (picture 4). Selecting the wrong picture is recorded as a wrong response.

Items are grouped in sets of 12 and all 12 items are presented in the given order within each set. The BPVS3 comprises a total of 14 sets (168 items) arranged in order of increasing difficulty. However, only an expected four to five sets (48 – 60 items) are administered to each child according to the following procedure. The test's administrator begins with the set indicated as an age-appropriate starting point for the child. The starting set for children aged 3 is set number 1 and the starting set for children aged 4 is set number 2. If the child makes more than one mistake in the starting set, the administrator moves to the lower set until they identify a so-called Basal Set (i.e., the lowest set for which 11 or 12 responses are correct).

They proceed to test forwards in sets until a so-called Ceiling Set is reached (i.e., the lowest set in which eight or more responses are wrong). Testing ends once all items in the Ceiling Set are tested. The administrator sums the number of wrong responses made across sets, from the Basal to the Ceiling Set included. Raw scores for each child are then calculated by subtracting the number of wrong responses from the number of the last item in the Ceiling Set. Raw scores range from 0 to 168.

Raw scores for each child are associated with age-standardised scores indicating the degree to which the child's raw score deviates from the average score in a nationally representative sample of children of the same age. The standardisation scale is based on a normal distribution with a mean standardised score of 100 and a standard deviation of 15, such that 68% of children of a certain age will score between 85 and 115 (i.e., within one standard deviation of the mean). Standardised scores higher than 140 or lower than 70 cannot be calculated with greater precision due to the low number of people who have scored this low or high in the standardisation sample. It is therefore recommended to record them as scores of 140+ and 70-, or as 69 and 141 if a specific number is needed.

The BPVS3 will be collected at four different time points - at baseline, in November 2023 for Cohort 1 and in November 2024 for Cohort 2, and at endline, in June-July 2024 for Cohort 1 and in June-July 2025 for Cohort 2.

The BPVS3 age standardised score will be used to compare children's language ability against a large nationally representative sample of children of similar age. Conversion tables used to standardise the raw BPVS scores are available in the BPVS3 Manual.

NatCen will collect assessment data on children by sending external assessors to visit the settings. These will be registered SLTs who will conduct baseline and endline assessments with children directly, one-to-one.

The same outcome will also be used as a measure of baseline attainment.

BPVS3 has several key advantages as a measure. Unlike alternative measures, such as the Preschool Early Literacy Indicator (PELI) used in the previous REAL evaluation, BPVS3 has been validated in the UK. It is quick to administer and score. It also has a high construct validity[3], meaning that it is considered a reliable measure of language ability. Moreover, BPVS3 was used in Hanen LLLI (Bury et al., 2022) and administered by SLTs with the same age group, with feedback indicating that they considered it age-appropriate and easy to administer.

Secondary outcome

As BPVS3 only measures receptive language, the Making it REAL evaluation will also use the Renfrew Action Picture Test (RAPT[4]) as a measure of receptive and expressive language. RAPT tests children's receptive and expressive language by asking them to describe pictures they are shown. The RAPT comprises 10 action picture cards depicting various scenarios. Children are shown one image at a time and asked a prompt question to describe the scenario. Their answers are recorded and scored along two dimensions - information and grammar -

[3] For more information, see: https://educationendowmentfoundation.org.uk/measures-database/british-picture-vocabulary-scale
[4] The Renfrew Action Picture Test | Expressive language skills (ndcs.org.uk)

according to question-specific criteria described in the scoring guides for each question. For example, an item would consist of a colour picture card illustrating a girl hugging a teddy bear, and the child would be asked what the girl is doing. Their answer would be scored for information (e.g., use of a "hug"-related verb, and a "teddy bear"-related noun) and for grammar (e.g., use of a relevant verb with and "-ing" ending), with each listed criterion contributing one or two points to be allocated to the response. Scoring criteria broadly increase in number and complexity with the order of items, such that the total number of points that can be allocated to a response is highest for item number 10. Information and grammar scores for each item are then summed separately to produce two overall scores per child: an information score and a grammar score. The raw score for information ranges from 0 to 41, while the raw score for grammar ranges from 0 to 39.

The main advantage of the RAPT is that it provides a snapshot of a child's level of expressive language. We will use the raw RAPT scores.

Baseline measures

BPVS3 will be used as a baseline measure obtained in November 2023 for Cohort 1 and November 2024 for Cohort 2 before children have access to the programme.

Data collection

Settings will provide details of children eligible to receive the intervention in October 2023 for Cohort 1 and October 2024 for Cohort 2 using NatCen's secure upload platform. This will include each child's first and last names, date of birth, Unique Pupil Number (UPN – if available), gender and eligibility for Pupil Premium funding. As children in PVI settings may not always have a UPN, we will collect home postcode to support NPD linkage in the future. This will allow a longer-term follow up. We will also use these children lists for baseline assessments (in November 2023 and November 2024) and endline assessments (in June-July 2024 and June-July 2025). Data will be collected from the same children at baseline and endline for BPVS and at endline only for RAPT.

Children's assessments will be conducted by trained SLTs on a one-to-one basis, recruited for the trial to minimise additional burden on setting staff and attrition due to the latter. This will also enhance measurement quality as SLTs will be blind to children's treatment allocation, unlike setting staff who will be aware of children's treatment allocation at endline.

Sample size calculations

Considering restrictions of the maximum delivery capacity for the Making it REAL programme, the evaluation will aim to recruit 150 settings, with half randomly allocated to the intervention condition and the other half randomly allocated to the usual practice control condition. Within each setting, six children will be selected to complete assessments.

Error! Reference source not found. shows the estimated Minimum Detectable Effect Size (MDES) for the trial without accounting for attrition at setting- and child-level. Our power calculations are informed by previous evaluations of REAL and Hanen LLLI. We assume there is a high child-level correlation between baseline and follow-up (0.70)[5] and moderate setting-

[5] Camilleri and Law (2014) provide the correlation between measurements of the BVPS taken 6-months apart, which was 0.90 (Table 1, page 511). We assume a more conservative correlation of 0.70.

level correlation (0.36) [6]. Power calculations were conducted using PowerUp! (Dong and Maynard, 2013).

Under these assumptions and an ideal scenario of no attrition, the trial would have power to detect an overall MDES of 0.195. The MDES for the EYPP would be 0.264.

Table 2: Sample size calculations (no attrition)

		Overall	EYPP
Minimum Detectable Effect Size (MDES)		0.195	0.264
Pre-test / post-test correlations	level 1 (child)	0.70	0.70
	level 2 (setting)	0.36	0.36
Intracluster correlations (ICCs)	level 2 (setting)	0.12	0.12
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		2	2
Average cluster size		6	2
Number of settings	Intervention	75	75
	Control	75	75
	Total	150	150
Number of children	Intervention	450	150
	Control	450	150
	Total	900	300

Notes: Power calculations were performed using PowerUp! under an alpha level of 0.05 and power of 0.8. The calculations include estimates of the proportion of variance explained through the included covariates at each of these levels (also known as R^2). The R^2 values here have been estimated by squaring the pre-test post-test correlation. An R^2 value of 0.49 at child-level and 0.13 at setting-level is used in the power calculations in PowerUp!. We note that EEF protocols usually include sub-group analysis by Free School Meal (FSM) status. Here we use EYPP rather than FSM, as EYPP data is directly available from settings, whereas FSM is not. EYPP provides schools with additional funding for all 3-year-olds from low-income families.

Attrition is a risk for any panel study and is a key consideration for this trial. In the previous REAL evaluation funded by EEF, 10 settings of 53 withdrew from the trial due to lack of staff capacity and COVID-19, giving an attrition of 19% (Speight et al., 2022). We do not expect such high attrition for Making it REAL. EEF will offer financial incentives to all settings to

[6] Bloom et al. (2007) estimate that R-squared at school level ranges from 0.18 to 0.73 for covariates measuring attainment years prior to the intervention. We assume a conservative case of R-squared equal to 0.13 (which translates to a setting-level correlation of 0.36). If this is higher, the study will then have more power and require a smaller sample.

promote engagement in the trial and minimise attrition, as described in the Design section above.

Although we do not expect high attrition due to the COVID-19 pandemic, the inclusion of PVI settings brings additional risks of higher setting- and child-level attrition, based on evidence from similar trials.[7] We have deliberately powered the trial to withstand a reasonable degree of attrition (i.e., 10% at setting and child level), based on learning from the previous REAL trial and our understanding of the potential risks of including PVI settings.

Given our expectations about attrition, Table 3 **Error! Reference source not found.**is based on the assumption that 135 settings will remain in the trial at endline (representing roughly 10% attrition across recruited settings), with an average of five children per setting (representing roughly 10% attrition across children). Under these assumptions the trial would still have power to detect a MDES of 0.21. To contextualise, Nutbrown et al. (2005, chapter 11, page 4) point out that an individual-level RCT of this programme found a positive impact with an effect size of 0.40 standard deviations on early literacy.

However, it should be noted that this MDES is less favourable than the 0.20 that would be required for the trial to achieve a '5-padlock' security rating. It is possible that we will be able to improve on this if pre-post correlations are higher than anticipated or if we are able to minimise attrition so that it is smaller than our current assumptions. In light of the recruitment challenges faced by the delivery team during the previous trial of REAL, we propose to work closely with them to refine the recruitment strategy in accordance with EEF guidance (EEF, 2022a).

We have also conducted power calculations to assess the MDES for subgroup analysis according to children's EYPP eligibility status. According to DfE (DfE, 2021a), 10% of pupils aged 3-4 are in receipt of EYPP across England. However, we expect that a higher percentage of children taking part in Making it REAL will be eligible for EYPP. This is because settings will be recruited from deprived areas and will be targeted based on the high proportion of low-income parents using them. Moreover, children eligible for EYPP are more likely to need support with literacy development (Roberts et al., 2017), so they will be more likely to be selected to participate in Making it REAL compared to children who are not eligible for EYPP. Through this lens, in our power calculations, we assume two EYPP children per setting. Assuming 10% setting and 10% child level attrition, the MDES for EYPP will be 0.289 (see Table 3). Subgroup impact analysis is likely to be underpowered given the proposed sample size and the level of attrition assumed. Interpretation of the results should be limited as there is a slightly higher chance that we will miss a statistically significant effect if the true effect size is less than 0.289. However, we should be able to reliably test if this effect size is significant or not.

Table 3: Sample size calculations (realistic scenario: 10% setting and 10% child level attrition)

	Overall	**EYPP**

[7] For example, the Maths Champions trial conduced in PVI settings encountered high levels of attrition at both setting and child level (Robinson-Smith et al., 2018). It is not clear why PVI nurseries appear to be more likely to drop out compared to school-based settings but the difference in pupil-level attrition seems to be largely due to children in PVI settings being more likely to move to a different setting (within the same school year) compared to children in school-based nurseries.

Minimum Detectable Effect Size (MDES)		0.210	0.289
Pre-test / post-test correlations	level 1 (child)	0.70	0.70
	level 2 (setting)	0.36	0.36
Intracluster correlations (ICCs)	level 2 (setting)	0.12	0.12
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		2	2
Average cluster size		5	2
Number of settings	Intervention	68	68
	Control	67	67
	Total	135	135
Number of children	Intervention	365	122
	Control	364	121
	Total	729	243

Notes: Power calculations were performed using PowerUp! under an alpha level of 0.05 and power of 0.8. The calculations include estimates of the proportion of variance explained through the included covariates at each of these levels (also known as R^2). The R^2 values here have been estimated by squaring the pre-test post-test correlation. An R^2 value of 0.49 at child-level and 0.13 at setting-level is used in the power calculations in PowerUp!. The average cluster size value is 5.4 and 1.8 in the two columns, accounting for 10% attrition at the pupil level. Numbers are rounded to the value of five and two in the table respectively. We note that EEF protocols usually include sub-group analysis by Free School Meal (FSM) status. Here we use EYPP rather than FSM, as EYPP data is directly available from settings, whereas FSM is not. EYPP provides schools with additional funding for all 3-year-olds from low-income families.

In our calculations, we also provide estimates for a more conservative attrition scenario of 20% at setting and 10% at child level (see Table 4). This study will be able to detect a MDES of 0.223 for the primary outcome analysis and 0.307 for the subgroup analysis.

Table 4: Sample size calculations (conservative scenario: 20% setting and 10% child level attrition)

		Overall	**EYPP**
Minimum Detectable Effect Size (MDES)		0.223	0.307
Pre-test / post-test correlations	level 1 (child)	0.70	0.70
	level 2 (setting)	0.36	0.36
Intracluster correlations (ICCs)	level 2 (setting)	0.12	0.12
Alpha		0.05	0.05
Power		0.8	0.8

One-sided or two-sided?		2	2
Average cluster size		5	2
Number of settings	Intervention	60	60
	Control	60	60
	Total	120	120
Number of children	Intervention	324	108
	Control	324	108
	Total	648	216

Notes: Power calculations were performed using PowerUp! under an alpha level of 0.05 and power of 0.8. The calculations include estimates of the proportion of variance explained through the included covariates at each of these levels (also known as R^2). The R^2 values here have been estimated by squaring the pre-test post-test correlation. An R^2 value of 0.49 at child-level and 0.13 at setting-level is used in the power calculations in PowerUp!. The average cluster size value is 5.4 and 1.8 in the two columns, accounting for 10% attrition at the pupil level. Numbers are rounded to the value of five and two in the table respectively. We note that EEF protocols usually include sub-group analysis by Free School Meal (FSM) status. Here we use EYPP rather than FSM, as EYPP data is directly available from settings, whereas FSM is not. EYPP provides schools with additional funding for all 3-year-olds from low-income families.

Randomisation

The unit of randomisation in this trial is at the setting-level. Settings that agree to participate will be stratified by local authority and setting type (school-based vs PVI setting) and randomised into a treatment or a control arm within each stratum. All participating settings will have a 50:50 chance of being assigned to the treatment or control group within each stratum.

Stratification will ensure that settings from the same region as well as settings with similar type of provision (i.e., school-based or PVI setting) will be evenly allocated to the treatment and control group. In practice, stratification helps us control for potential variations in setting characteristics and programme implementation across geographical areas, and thus decreases the variance and improves the precision of the impact estimator.

Our plan is to randomise settings after baseline data collection (i.e., in November) to minimise the risk of attrition and to facilitate preparations for training delivery undertaken by the developer. The 75 Cohort 1 recruited settings will be randomised in November 2023, while the 75 Cohort 2 recruited settings in November 2024. NatCen will inform Cohort 1 and Cohort 2 settings about their allocation in November 2023 and November 2024, respectively, once baseline data collection is complete. Settings randomised into control for Cohort 1 will be eligible to sign up and be randomised again in year 2, alongside newly recruited settings.

Randomisation will be undertaken in Stata and both *.do* and *.log* files will be used to record the randomisation process. One consideration when conducting stratified randomisation is having unequal numbers of settings from some areas or types of settings. This could mean that there could be a high probability that the treatment or control group will be of unequal size. To account for unequal allocation, we will use the *randtreat* command and the option *misfits(global)* in Stata. This command also allows us to set a seed (we will use a random number from random.org) so that the random assignment can be replicated.

At the time of randomisation, the researchers will be blinded to settings' identity. To achieve this, setting identifiers will be removed prior to randomisation from the dataset. This way the

researcher will not be able to see which setting is allocated to which condition. Setting identifiers will be merged with group allocation data after randomisation.

Analysis

Primary analysis

We will use an Intention-to-Treat (ITT) approach to estimate the impact of Making it REAL programme on the primary outcome. ITT will mean that all those allocated to intervention and control conditions in the randomisation will be included in the final analysis as per their original group assignment, so long as they provide data. The trial is designed as a two-armed cluster randomised efficacy trial with children (level one) clustered within settings (level two). To account for the clustering of children within settings, the impact will be estimated using a two-level linear regression model.

The BPVS3 age-standardised score at follow-up will be the dependent variable, with a binary indicator of treatment allocation, baseline BPVS3 (age-standardised) score and randomisation stratifier variables (local authority and setting type) included as independent variables. Setting-level random effects will be included in the model by allowing the intercept to vary by setting. Our model will follow EEF statistical analysis guidance (EEF, 2022b).

Secondary analysis

We will estimate the impact of Making it REAL programme on the secondary outcome (RAPT). The analytical approach will be analogous to the primary outcome estimation, consisting of a two-level linear regression model with children nested within settings. This analysis would use the same specification as the primary analysis, with the RAPT score at follow-up as the dependent variable[8]. We will additionally control for age since we will be using raw (non-standardised) RAPT scores as outlined above (see Secondary outcome

Subgroup analysis

Subgroup analysis will be undertaken to explore if Making it REAL has a differential impact for children who are eligible for the Early Years Pupil Premium. As children who are registered as eligible for EYPP are considered a key target group by the EEF, we will assess whether the treatment effect varies between EYPP children and non-EYPP children.

To estimate the subgroup analysis, we will use interaction terms between the treatment indicator and EYPP. Yet, it is worth highlighting that this analysis is likely to be underpowered given the proposed sample sizes, so interpretation of results may be limited.

Effect size

Effects will be presented as *Hedges' g* effect sizes with 95 percent confidence intervals, accounting for clustering of children within settings (Hedges, 2007).

Compliance

[8] RAPT analysis would use the BPVS3 baseline measure as a covariate (since we will not be capturing the RAPT itself at baseline).

Compliance is defined as the fulfilment of a set of minimum criteria which determine whether a setting has delivered the Making it REAL programme as intended. Full compliance is defined for when a child has received all four home visits and their parents have attended at least two literacy events (out of four).

Specifically, the four home visits should be:

- at least 40 mins long
- spaced throughout the school year, with two home visits in the spring term and two visits in the summer term; and
- at each visit both the child and a parent/primary caregiver would need to be present.

Data on home visits and attendance at the literacy events will be captured via (compliance spreadsheet) templates designed by NatCen and completed by Making it REAL practitioners.

In terms of analysis, compliance will be a binary measure, indicating whether a child was fully compliant or not. In a situation of imperfect compliance (e.g., children receiving fewer than four home visits or parents attending fewer than two literacy events), we will undertake a Complier Average Causal Effect (CACE) analysis, by drawing on an instrumental variable (IV) approach and using a two-stage least squares (2SLS) estimation approach to obtain the treatment effect for those who complied with assignment (see Imbens and Angrist, 1994).

The first stage of the IV estimation estimates whether assignment to the Making it REAL programme pushes children to take up treatment (the first stage regresses treatment assignment on compliance). This provides an estimate of the compliance rate. Results for the first stage will report the correlation between the instrument and the endogenous variable as well as an F-test (EEF, 2022b). The second stage of the IV estimation predicts the outcome using the compliance rate estimated in the first regression by substituting the treatment indicator (i.e., assignment to Making it REAL programme) with the compliance rate. The results of this model will answer the research question: ‘To what extent does compliance with the Making it REAL delivery requirements lead to improved language outcomes for children?’. We will only use the model for estimations for the primary outcome measure, i.e., for receptive language skills.

Missing data analysis

We will follow EEF’s statistical analysis guidance to address missing data (EEF, 2022b), the selected option will be decided after we have a better understanding of the extent of missingness and the patterns of missingness. A detailed account of our missing data approach will be outlined in the Statistical Analysis Plan (SAP).

Implementation and process evaluation (IPE) design

IPE dimensions

The IPE will cover the following dimensions:

- **Fidelity** – extent to which settings deliver Making it REAL as intended

- **Dosage** – how many home visits and literacy events settings deliver, when they are delivered and how long they last
- **Quality** – quality of intervention activities (practitioner training, network events, home visits and literacy events)
- **Adaptation** – extent to which practitioners vary in how they approach home visits and literacy events, and what the variations are
- **Reach** – extent to which Making it REAL reaches its intended population
- **Responsiveness** – extent to which settings, practitioners, parents and children actively engage with the intervention
- **Programme differentiation** – how the intervention differs from usual practice
- **Perceived outcomes** – practitioners' and parents' perceptions of the effects of the programme on children, parents, and practitioners
- **Costs** – the cost of delivering the programme in settings
- **Context/moderators** – factors that influence the delivery or perceived outcomes of the intervention, e.g., socioeconomic background of families, setting type, practitioners' qualification level.

Research questions

The IPE will answer the following questions:

- RQ3 (IPE): How is the programme different to usual practice?
- RQ4 (IPE): To what extent is the programme delivered as intended?
- RQ5 (IPE): What are the facilitators and barriers to delivery?
- RQ5a (IPE): Are there particular facilitators and barriers for disadvantaged families, from being selected into the programme to their participation in the activities?
- RQ6 (IPE): What are the perceived outcomes of the programme on practitioners?
- RQ7 (IPE): What are the perceived outcomes of the programme on parents?
- RQ8 (IPE): What are the perceived outcomes of the programme on children?
- RQ9 (IPE): What are the effects of the programme on early literacy activities in the home learning environment?
- RQ10 (IPE): What is the per-child cost of the programme?

Research methods

NatCen will conduct a multi-stage implementation and process evaluation and take a mixed-method approach to data collection. This will include observations, in-depth interviews,

surveys, and collection of compliance and cost data. This approach to data collection will allow us to capture experiences and views of multiple practitioners and parents involved in the trial.

Below, we have mapped our research design against key IPE dimensions (Table 5).

Table 5: IPE methods and domains

Method	IPE domain									
	Fidelity	**Dosage**	**Quality**	**Adaptation**	**Reach**	**Responsiveness**	**Usual practice**	**Perceived outcomes**	**Costs**	**Context**
Training observations	■		■			■				
Network meeting observations			■	■		■				
Literacy event observations	■		■	■	■	■				
Practitioner interviews	■		■	■	■	■	■	■		■
Parent interviews						■	■	■		
Practitioner survey		■	■		■	■	■	■	■	■
Parent surveys						■	■	■		
Cost data									■	
Compliance data	■	■			■					

Note: shaded cells indicate which IPE domain is covered by which IPE method.

Training observations

NatCen will observe the two-day, in-person training events for practitioners in Cohort 1 and Cohort 2. We will prepare a detailed observation protocol to guide data collection during events. The protocol will focus on **fidelity** (is training content aligned with the intended model, RQ4); **quality** (how well training is delivered) and **responsiveness** (how well practitioners respond to content and delivery). The observation protocol will also aim to capture any differences between the training sessions for Cohorts 1 and 2. Observations will take place in December 2023 and December 2024.

Network meeting observations

NatCen will also observe four network meetings delivered by NCB to practitioners in the treatment group, two in Cohort 1 and two in Cohort 2. If these events are delivered in different formats (in person and online), we will aim to observe both types of events evenly. We found during our previous evaluation that these events focused on programme adaptations. Our observation guide will therefore focus on **adaptation** (what changes do practitioners make to the activities), **quality** (how well the meetings are delivered) as well as **responsiveness** (how well practitioners respond to the content and delivery of these meetings). We will also capture any differences between the network meetings for Cohorts 1 and 2. Observations will take place between January-June 2024 and January-June 2025 and last 60 minutes each.

Literacy event observations

NatCen will observe four group literacy events that practitioners will deliver to parents, two in Cohort 1 and two in Cohort 2. We will aim to include a mix of PVI and school-based settings in the sample. The protocol will focus on **fidelity** (are events aligned with the intended model, RQ4), **quality** (are events delivered well), **adaptation** (are changes are made to the intended model for literacy events), **reach** (how many parents attend events) and **responsiveness** (how well parents engage with the content and delivery). Observations will take place in February-June 2024 and 2025 and will last 60 minutes each.

Practitioner interviews

NatCen will interview 16 practitioners in the treatment group, eight in Cohort 1 and eight in Cohort 2. Practitioners will be recruited from the sample frame NCB provide. We will include a mix of settings in the sample. We will develop a topic guide to explore:

- **fidelity** (to what extent do activities match with the intended model),
- **quality** (quality of interactions between practitioners and families during home visits and literacy events; quality of training and network events),
- **adaptation** (what changes do practitioners make to the programme and why),
- **reach** (to what extent do families participate in activities and are there differences based on social disadvantage),
- **responsiveness** (to what extent do parents, children and practitioners engage with activities),

- **usual practice** (what parent engagement activities did practitioners do pre-intervention; RQ3),
- **perceived outcomes** on practitioners, parents and children, and,
- **context/moderators** (who did the activities work for and under what circumstances).

Regarding perceived outcomes on children, we will focus on four strands of literacy and ask about: children's understanding of environmental print, children's interest in books, children's early writing skills and children's oral language. Of these child outcomes, only one – children's oral language – is being measured in the impact evaluation, but the IPE will cover all four.

We will conduct qualitative interviews online or by telephone and they will last up to 60 minutes. Interviews will take place after delivery has completed in June-July 2024 and June-July 2025. We will only interview practitioners at endline to reduce burden on practitioners. Practitioners who take part in an interview will be offered a £30 online gift card as a thank-you.

Parent interviews

NatCen will interview 16 parents who took part in the intervention, eight in Cohort 1 and eight in Cohort 2. We will recruit parents through practitioners, which was a successful method in our previous evaluation. We will include parents from different types of settings. The topic guide will explore **responsiveness** (to what extent do children and parents engage with activities) and **usual practice** (what home learning activities do parents normally do with children, and to what extent are REAL activities different; RQ3). We will also ask about **perceived outcomes** on parents (RQ7) and children (RQ8).

Qualitative interviews will last 30 minutes, and they will be conducted on the phone. The parent interviews will take place at endline in June-July 2024 and June-July 2025. Parents who participate in an interview will be offered a £20 online gift card as a thank-you for taking part.

Practitioner survey

NatCen will conduct two online surveys with practitioners: an endline survey in Cohort 1 and in Cohort 2. We will send the survey link to all practitioners as a census survey. The survey will collect mainly quantitative data on **dosage** (how many home visits and literacy events practitioners delivered, to supplement compliance data), **quality** (perceptions of quality of training and network events; whether practitioners used the ORIM framework when planning home visits, whether practitioners focused on all four strands of early literacy), **reach** (to what extent families participated in activities and whether there were differences based on social disadvantage) and **responsiveness** (to what extent practitioners and the families engaged with activities). The questionnaire will also collect quantitative data on **perceived outcomes** on practitioners (RQ6), parents (RQ7) and children (RQ8) and **usual practice** (RQ3). Lastly, will ask about **costs**. Data on the type of setting practitioners were based at, their level of qualifications and years of experience will be used to capture **context/moderators** (to explore whether type of setting and characteristics of practitioners were important moderators for how the intervention was delivered and experienced).

Practitioner surveys will take between 10-15 minutes to complete. They will take place at endline in June-July 2024 and June-July 2025. We will not carry out a baseline practitioner survey to reduce burden on practitioners during the busy baseline testing period.

Parent surveys

NatCen will conduct four online surveys with parents taking part in the evaluation: at baseline and endline in Cohort 1 and Cohort 2. Parents in both intervention and control group settings will be invited to take part. We will ask practitioners to share with parents a letter (hard copy) with a personalised link to the baseline survey. The baseline survey will collect parents' contact details (name and email address) so that we can contact them at endline directly.

The main aim of the parent survey is to collect data about early literacy activities in the home learning environment (HLE). Rather than using available measures of HLE, we will design a module of questions that is tailored to Making it REAL. This module of questions will allow us to measure the effects of the programme on the HLE (RQ9) and to capture **usual practice** (RQ3). At endline, we will also ask parents in the intervention group questions about their experience of taking part in Making it REAL. This will allow us to assess **responsiveness** (how well they and their children engage with activities) and **perceived programme outcomes** on parents (RQ7) and children (RQ8).

Surveys will take up to 10 minutes to complete. Baseline surveys will take place in November 2023 and November 2024. Endline surveys will take place in June-July 2024 and June-July 2025. We will offer all parents who complete the survey an incentive, both at baseline (a £5 online gift card) and endline (a £10 online gift card).

Compliance data

NatCen will collect data from practitioners in the intervention group about home visits and literacy events. For home visits, they will be asked to record for each child taking part in the intervention how many home visits they had, dates of those visits and their length. For literacy events, practitioners will be asked to record the number and dates of the events and whether they were attended by the families taking part in the intervention. This data relates to the IPE dimensions of **fidelity, reach** and **dosage**.

In the previous REAL trial funded by the EEF, NCB collected and quality-assured data before securely sharing it with NatCen; we will follow the same process in the Making it REAL evaluation. NatCen will ask for the compliance data on a termly basis as this aligns with how practitioners plan activities. Practitioners will be asked to complete compliance spreadsheets twice: one for the spring term and one for the summer term.

In the previous REAL trial, incomplete or inaccurate compliance data made it challenging to analyse the results. In this evaluation, NatCen will simplify the spreadsheet design and ask fewer questions to make the sheet easier for practitioners to navigate. We will also pre-fill the spreadsheets with serial numbers of the participating children to avoid errors in serial number entries. Lastly, we will be set the compliance spreadsheets at the setting-level rather than practitioner-level, with the expectation that practitioners will coordinate with each other to complete data for all the children in their setting.

Completing the spreadsheets will take practitioners a total of about 30 minutes (15 minutes in the spring term and 15 minutes in the summer term).

Table 6 provides an overview of the number of participants and encounters for each IPE activity.

Table 6: IPE methods overview

Research method	Data collection methods	Number of encounters and participants	Data analysis methods	Research questions addressed	IPE dimension
Training observations	In person	Encounters: 2 observations Participants per encounter: 5+	Qualitative	4, 5	Fidelity, quality, responsiveness
Network meeting observations	In person	Encounters: 4 observations Participants per encounter: 5+	Qualitative	5	Quality, adaptation, responsiveness
Literacy event observation	In person	Encounters: 4 observations Participants per encounter: 5+	Qualitative	4, 5	Fidelity, quality, adaptation, reach, responsiveness
Parent survey	Online	Encounter: 4 surveys Participants per encounter: 450	Quantitative	3, 5, 7, 8, 9	Usual practice, responsiveness, perceived outcomes
Practitioner survey	Online	Encounters: 2 surveys Participants per encounter: 75	Quantitative	3, 5, 5a, 6, 7, 8	Dosage, reach, responsiveness, perceived outcomes, usual practice, costs, quality, context/moderators
Practitioner interview	Online/Telephone	Encounters: 16 interviews Participants per encounter: 1	Qualitative	3, 4, 5, 5a, 6, 7, 8	Fidelity, quality, adaptation, reach, responsiveness, usual practice, perceived outcomes, context/moderators
Parent interview	Telephone	Encounter: 16 interviews Participants per encounter: 1	Qualitative	3, 5, 7, 8	Usual practice, responsiveness, perceived outcomes
Compliance spreadsheets	Online	Encounters: 4 (termly) Participants per encounter: 37-38 settings	Quantitative	4	Fidelity, dosage and reach
Costs proforma	Online	Encounter: 1 (NCB)	Quantitative	10	Costs

Analysis

NatCen will digitally record (with participant permission) all interviews, and they will be professionally transcribed. All observations will be recorded using detailed fieldnotes.

We will manage and analyse interview data using the Framework approach, developed by NatCen (Ritchie et al, 2014). Using this approach, we will develop thematic analytical frameworks for each research activity, using themes from the topic and observation guides and other themes emerging from the data. Each framework will be assembled into a matrix, where each row will represent an individual interview or observation and each column a theme and any related sub-themes. This approach allows analysis within and across cases and themes. Transcripts and observation notes will be reviewed in detail, and data will be summarised and categorised systematically by theme, whilst using illustrative verbatim quotes where appropriate.

Once all data has been included in the Framework matrix, we will move onto analysis. Our analysis approach will involve a phase of 'detection', including studying elements of participants responses about a particular phenomenon, listing these and then sorting them thematically. Once we have identified different themes in the data, we will create higher-level categories that work as meaningful conceptual groupings for participants' views and experiences within and across settings.

We will manage and analyse survey data using SPSS. This data will be analysed using descriptive statistics and cross-tabulations. Where possible, we will look at differences in parental practice pre- and post-intervention and between control and treatment groups.

We will triangulate and synthesis the IPE data according to our research questions, implementation dimensions and logic model components. This will provide a comprehensive assessment of implementation, help to explain findings from the impact evaluation and provide useful lessons for future delivery.

Cost evaluation design

NatCen will collect cost data using an approach recommended in the EEF guidelines, called 'the ingredient method' (EEF, 2023). Using this approach, we will account for all costs incurred to deliver the programme regardless of who incurs the costs. This includes costs of resources, materials, personnel time to deliver and attend training, personnel time to implement the intervention, expenses and incentives. We will collect cost data using the following methods:

1) Cost pro-forma sent to the delivery team at NCB will be used to calculate the cost of delivering training, network events and any support for settings.
2) The endline practitioner survey will collect data on staff time to attend training and network events and to deliver intervention activities, including information on whether settings used supply staff for cover.

During analysis, we will divide programme costs into pre-requisite costs, set-up costs and recurring costs. The cost evaluation will estimate average marginal costs per pupil per year over a three-year period.

Ethics and registration

NatCen has a robust ethics procedure. NatCen's Research Ethics Committee (REC) reviewed and approved NatCen's research proposal for this project on 31 January 2023. The REC procedure is designed to provide ethical guidance and advice, and to ensure all research undertaken by NatCen meets the ethical standards of government and other funders. NatCen's ethics procedure is aligned with the 2021 Government Social Research (GSR) professional guidance and the 2021 Social Research Association (SRA) ethics guidance.

Agreement to participate in the trial

Recruitment of settings

NCB are responsible for recruiting settings into the trial. The delivery team will work with around five local authorities in Yorkshire and the Humber. They will support NCB with recruiting school-based settings and those in the PVI section. As part of the recruitment campaign, NCB will hold information events for practitioners in each local authority.

NCB began recruitment in April 2023 for Cohort 1. NatCen drafted the trial recruitment documents, including a setting information sheet, a memorandum of understanding (MoU), a parent information sheet, a Taking Part webpage and a privacy notice.

The setting information sheet includes a brief description of the Making it REAL programme, what taking part in the programme and evaluation involves, the benefits of taking part, and eligibility criteria. The MoU explains in more detail the expectations and responsibilities for settings in the treatment and the control groups. By signing the MoU, settings will formally agree to take part in the evaluation and to the conditions associated with taking part.

Recruitment of practitioners

Two practitioners at each setting will deliver the intervention (if randomised to treatment). In signing the MoU, settings will agree to share names and email addresses of these practitioners

(e.g., teachers/teaching assistants in schools and early years educators in PVI early years settings) with NatCen. NCB will securely share with NatCen a list of settings and contact details for the practitioners selected to deliver Making it REAL in their setting. NatCen will use these contact details to send personalised invites to practitioners to take part in interviews and a survey. For interviews, practitioners will be sampled to take part from different settings, to reduce the burden on an individual setting.

Recruitment of parents

Impact evaluation

Practitioners will shortlist four children each (eight in total per setting) to potentially take part in Making it REAL. They will make selections based on eligibility criteria shared by NCB (this will be shared with settings at a later stage; however, the programme will be aimed at those children most likely to not meet a good level of development at the Early Years Foundation Stage Profile). Practitioners will hand out information sheets (drafted by NatCen) to the parents (or carers or guardians) of children selected to take part.

The parent information sheets will explain that their child's nursery setting will securely share information about their child with NatCen, and that their child will take part in baseline and endline assessments. The information sheet will also inform parents of their right to opt their child out of the evaluation and the Making it REAL intervention, and that doing so means their children will not take part in any testing nor will their information be shared.

While practitioners will shortlist eight children in each setting, only six will take part in the Making it REAL evaluation. Nurseries will be asked to shortlist eight children in case some children are absent on the day of the baseline assessments, or in case of late parental opt-outs. The shortlist will help keep the trial sufficiently powered from the start. If all the children from the shortlist are present on the day of the baseline assessments, the evaluation team at NatCen will have randomly selected six children to take part in the study, and two children to not take part. The parent information sheet explains the shortlist to parents.

Parent surveys

NatCen will invite all parents in treatment and control group settings to take part in two online surveys, one at baseline and the other one at endline. For the first survey, we will ask practitioners to share a letter with parents in person. The letters will have a personalised online survey link. The baseline survey will then collect parents' contact details so that NatCen can contact parents directly to take part in the endline survey at the end of the school year.

Parent interviews

NatCen will request practitioners to ask parents who took part in the Making it REAL intervention if they would like to take part in an interview. We will provide practitioners with an information sheet about the interviews to share with parents. The information sheet will explain what the interview will be about, that participation is voluntary and how confidentiality and anonymity will be managed. If parents are happy to share their contact details with the research team, the practitioners will pass along their phone number and email address. A member of the evaluation team will get in touch with parents to schedule an interview.

Trial registration

The evaluation has been registered on the Open Science Framework (OSF) platform and can be accessed at the following link: https://osf.io/xj7dg.

Data protection[9]

During the evaluation

For the duration of this evaluation, NatCen is the data controller and data processer. Our legal basis for processing the data is 'legitimate interest'. This means that we believe that there is a genuine reason for us to process this data (to evaluate the impact of Making it REAL), that this data is needed to fulfil this purpose (we could not evaluate it without this information), and that using this data will not interfere with individuals' interests, rights or freedoms.

NatCen will receive personal data from the delivery team at NCB, including names and contact details for staff working at the delivery settings. Data sharing between NatCen and NCB is governed by a data sharing agreement (DSA).

NatCen will receive personal data relating to the children directly from settings through a secure website. For the IPE, we will collect parents' names and contact details through an online survey as well as through practitioners with the parents' consent.

NatCen will collect primary data during the evaluation: 1) we will use data from the language-based assessments to inform the impact evaluation, and 2) we will use data collected from interviews, surveys, observations, compliance spreadsheets and a cost pro-forma to inform the IPE. The data subjects will be staff working at the delivery settings, 3 and 4-year-old children and their parents taking part in Making it REAL as well as the delivery team at NCB. No special category data will be collected as part of the evaluation.

NatCen will store and handle all data securely and confidentially in line with UK GDPR. Only named individuals at NatCen will have access to data collected as part of the evaluation. This will be monitored through a data security plan set up by the Principal Investigator. In addition, some third parties (e.g., transcribers, printers and SLTs) will have limited access to personal data. You can read more about who has access to data collected as part of this evaluation on NatCen's privacy notice for Making it REAL.

After the evaluation

Reports and publications arising from this research will not identify any individual research participant or setting. All personal information, and any other data held on the project, will be securely deleted from NatCen systems within twelve months of the project's completion, in March 2027.

NatCen will transfer the impact evaluation data to the EEF data archive held within the Office for National Statistics Secure Research Service (ONS SRS). The EEF are the data controller for the data archive which is managed by FFT Education who act as a data processor working on EEF's behalf. NatCen will submit the data directly to FFT through a secure portal. In addition to impact evaluation data, NatCen will also transfer to the archive personal

[9] Please see the Data Protection Statement for EEF Evaluations.

information about children taking part in the evaluation (i.e., child's Unique Pupil Number, full name, date of birth and home postcode). This is to enable potential follow-up research including linking of data from Making it REAL with data from the National Pupil Database and other datasets. For further information, see the EEF's archive privacy notice.

Personnel

A team of research and evaluation specialists at NatCen will carry out the evaluation. Centre for Children and Families will lead the evaluation with support from impact evaluation specialists from the Centre for Evaluation.

Table 7: Evaluation team

Delivery team	Institution	Role
Dr Svetlana Speight	Research Director, Centre for Children and Families	Principle investigator
Maha Basharat	Researcher Director, Centre for Evaluation	Impact evaluation lead
Helena Takala	Senior Researcher, Centre for Children and Families	IPE lead and day-to-day project manager
Dr Konstantinos Papaioannou	Senior Researcher, Centre for Evaluation	Impact evaluation support
Alina Fletcher	Researcher, Centre for Children and Families	IPE support
Ekaterina Stoilova	Researcher, Centre for Evaluation	Impact evaluation support
Dr Tina Haux	Director, Centre for Children and Families	IPE Quality Assurance
Dr Andi Fugard	Deputy Director, Centre for Evaluation	Impact evaluation Quality Assurance

Table 8: Delivery team

Delivery team	Institution	Role
Ellie Suggate-Francis	Assistant Director – Early Childhood Unit, National Children's Bureau (NCB)	Director of delivery
Laura Sarmiento	National Children's Bureau (NCB)	Programme manager
Octavia Holland	National Children's Bureau (NCB)	Strategic oversight
NCB Associates	National Children's Bureau (NCB)	Delivery of training and network events

Risks

Table 9: Risk management

Risk	Likelihood/ Impact	Mitigation/Contingency
Settings, practitioners or parents reluctant to sign up to trial	**Likelihood:** Medium **Impact**: Medium	Recruitment was successful for REAL. We will work closely with the developers and adapt previous materials allowing more time for recruitment. In the second year (Cohort 2), there may be an option to increase the number of settings recruited or modify the recruitment strategy if Cohort 1 experienced substantial recruitment problems.
Setting attrition	**Likelihood:** Medium **Impact:** High	MoUs will outline details of requirements for participating settings. Research design will minimise setting burden whilst ensuring depth of data. Power calculations account for potential additional attrition due to PVI inclusion.
Children's parents or settings oppose child testing	**Likelihood:** Medium **Impact:** High	MoUs will outline requirements for participating settings. Information for parents will use accessible language. A project web page will enable parents to contact the research team. We will ask settings to select two 'reserve' children to account for late parent opt-outs at baseline and to make sure six children are tested at baseline at every setting. Settings will be incentivised to support testing (£150 at baseline and £150 at endline, part of delivery budget)
Pupil- and practitioner-level attrition	**Likelihood:** Medium **Impact:** High	We account for attrition in our power calculations to anticipate potential high attrition. If a practitioner leaves an intervention group setting or is unable to continue with the intervention for another reason, we will work with settings to see if the remaining practitioner could complete remaining home visits with the selected children in their place.
Evaluation team member absence or turn-over	**Likelihood:** Medium **Impact:** Low	NatCen staff have a three-month notice period to allow time for handover. The team can be supplemented by researchers with relevant topic and methodological expertise from a pool of researchers.

Timeline

This section includes two timelines for the evaluation. The first one covers key delivery and evaluation activities for Cohort 1 (2023-24) and Cohort 2 (2024-25). The second timeline highlights key dates for analysis, outputs and data archiving.

Table 10: Timeline for delivery and evaluation activities

Dates	Activity	Responsibility
April-July 2023 for Cohort 1 and January-July 2024 for Cohort 2	Recruitment of settings	NCB
September	Recruitment of families	Settings, NCB
October	Child enumeration	Settings, NatCen
November	Baseline testing of children	NatCen
November	Randomisation of settings	NatCen
November	Parent baseline survey	NatCen
December	Practitioner training	NCB
December	Observations of practitioner training	NatCen
January-June	Programme delivery	Settings, NCB
February-June	Observations of network meetings and literacy events	NatCen
June-July	Parent interviews	NatCen
June-July	Endline testing of children	NatCen
June-July	Practitioner interviews	NatCen
June-July	Practitioner survey	NatCen
June-July	Parent endline survey	NatCen

Table 11: Timeline for analysis, outputs and data archiving

Dates	Activity	Responsibility
May 2023	Publication of evaluation protocol	NatCen, EEF
February 2025	Publication of Statistical Analysis Plan	NatCen, EEF
August-November 2025	Analysis and report writing	NatCen
January 2026	First draft report submitted	NatCen
April 2026	Data archiving	NatCen, EEF

References

Asmussen, K., Feinstein, L., Martin, J. and Chowdry, H., 2016. *Foundations for Life: What Works to Support Parent Child Interaction in the Early Years*. [pdf] Early Intervention Foundation. Available at: https://www.eif.org.uk/files/pdf/foundations-for-life.pdf.

Axford, N., Sonthalia, S., Wrigley, Z., Goodwin, A., Ohlson, C., Bjornstad, G., Barlow, J., Schrader-McMillan, A., Coad, J. and Toft, A., 2015. *The Best Start at Home: What Works to Improve the Quality of Parent-Child Interactions from Conception to Age 5 Years? A Rapid Review of Interventions*. [online] Early Intervention Foundation. Available at: https://www.eif.org.uk/report/the-best-start-at-home.

Bloom, H. S., Richburg-Hayes, L., Rebeck Black, S., 2007. Using covariates to improve precision for studies that randomize schools to evaluate educational interventions. *Educational Evaluation and Policy Analysis*, 29(1), 30-59. https://www.jstor.org/stable/30128044

Bury, J., Marshall, L., Read, H., Roberts, E., Fletcher, A., and Scandone, B. 2022. *Hanen Learning Language and Loving It (LLLI) Evaluation Report*. [pdf] Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/projects/Hanen-LLLI_report_finalised.pdf?v=1680093510.

Camilleri, B., & Law, J. 2014. Dynamic assessment of word learning skills of pre-school children with primary language impairment. *International journal of speech-language pathology*, *16*(5), 507-516. https://doi.org/10.3109/17549507.2013.847497

Department for Education, 2018. *Improving the home learning environment: A behaviour change approach*. [pdf] Department for Education. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/919363/Improving_the_home_learning_environment.pdf.

Department for Education, 2021a. *Education provision: children under 5 years of age*. [online] Available at: https://explore-education-statistics.service.gov.uk/find-statistics/education-provision-children-under-5/2021.

Department for Education, 2021b. *Statutory framework for the early years foundation stage*. [pdf] Department for Education. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974907/EYFS_framework_-_March_2021.pdf.

Department for Education, 2022a. *Early years foundation stage profile results*. [online] Available at: Early years foundation stage profile results, Academic year 2021/22 – Explore education statistics – GOV.UK (explore-education-statistics.service.gov.uk).

Department for Education, 2023. *Early years education recovery programme: supporting the sector*. [online] Available at: https://www.gov.uk/government/publications/early-years-education-recovery-programme/early-years-education-recovery-programme-supporting-the-sector#pdp3.

Dong, N. and Maynard, R., 2013. *PowerUp!*: A Tool for Calculating Minimum Detectable Effect Sizes and Minimum Required Sample Sizes for Experimental and Quasi-Experimental Design Studies. *Journal of Research on Educational Effectiveness*, [e-journal] 6(1), pp.24-67. https://doi.org/10.1080/19345747.2012.673143.

Education Endowment Foundation, 2022a. *Guidance on Recruitment and Retention*. [pdf] Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/Recruitment-and-retention-guidance_2022.pdf?v=1680078497.

Education Endowment Foundation, 2022b. *Statistical analysis guidance for EEF evaluations*. [pdf] Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/evaluation/evaluation-design/EEF-Analysis-Guidance-Website-Version-2022.14.11.pdf?v=1679395501.

Education Endowment Foundation, 2023. *Cost evaluation guidance for EEF evaluations*. [pdf] Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/evaluation/evaluation-design/Cost-Evaluation-Guidance-Feb_2023.pdf?v=1680078496.

Hannon, P., 1995. *Literacy, Home and School: Research and Practice in Teaching Literacy with Parents*, London: Falmer Press.

Hannon, P., Nutbrown, C. and Morgan, A., 2020. Effects of extending disadvantaged families' teaching of emergent literacy. *Research Papers in Education*, [e-journal] 35(3), pp.310-336. https://doi.org/10.1080/02671522.2019.1568531.

Hedges, L. V., 2007. Effect Sizes in Cluster-Randomized Designs. *Journal of Educational and Behavioral Statistics*, [e-journal] 32(4), pp.341-370. https://doi.org/10.3102/1076998606298043.

Hirst, K., Hannon, P. and Nutbrown, C., 2010. Effects of a preschool bilingual family literacy programme. *Journal of Early Childhood Literacy*, [e-journal] 10(2), pp.183-208. https://doi.org/10.1177/1468798410363838.

Imbens, G.W. and Angrist, J.D., 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica*, [online] Available at: https://www.jstor.org/stable/2951620?seq=9.

Law, J., Charlton, J., McKean, C., Beyer, F., Fernandez-Garcia, C., Mashayekhi, A. and Rush, R., 2018. *Parent-child reading to improve language development and school readiness: A systematic review and meta-analysis*. [pdf] London: Nuffield Foundation. Available at: https://www.nuffieldfoundation.org/wp-content/uploads/2019/11/B804D16A-F0A4-43C0-920F-4A274A131AAF.pdf.

Lehrl, S., Evangelou, M. and Sammons, P., 2020. The home learning environment and its role in shaping children's educational development. *School Effectiveness and School Improvement*, [e-journal] 31(1), pp.1-6. https://doi.org/10.1080/09243453.2020.1693487.

McNeal Jr, R. B., 2014. *Parent Involvement, Academic Achievement and the Role of Student Attitudes and Behaviours as Mediators*. [pdf] Horizon Research Publishing. Available at: https://files.eric.ed.gov/fulltext/EJ1053945.pdf.

Melhuish, E. and Gardiner, J., 2018. *Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to age four years*. [pdf] Department for Education. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1034449/SEED-Age_4_RESEARCH_REPORT.pdf.

Melhuish, E. and Gardiner, J., 2020. *Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to age five years*. [pdf] Department for Education. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/867140/SEED_AGE_5_REPORT_FEB.pdf.

Nutbrown, C. and Hannon, P., 2003. Children's Perspectives on Family Literacy: Methodological Issues, Findings and Implications for Practice. *Journal of Early Literacy Research*, [e-journal] 3(2), pp.115-145. https://doi.org/10.1177/14687984030032001.

Nutbrown, C., Hannon, P. and Morgan, A., 2005. *Early Literacy Work with Families: Policy, Practice and Research*. London: SAGE Publications Ltd.

Ritchie, J., Lewis, J., Nicholls, C.M. and Ormston, R. eds., 2014. Qualitative Research Practice: A Guide for Social Science Students and Researchers. 2nd ed. London: SAGE Publications Ltd.

Roberts, E., Griggs, J., Robb, S., 2017. *Study of Early Education and Development: Experiences of the Early Years Pupil Premium*. Department for Education. Available at: https://www.gov.uk/government/collections/study-of-early-education-and-development-seed

Robinson-Smith, L., Fairhurst, C., Stone, G. et al., 2018. *Maths Champions: Evaluation Report and Executive Summary*, Education Endowment Foundation. Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/maths-champions

Siraj-Blatchford, I. and Siraj-Blatchford, J., 2009. *Improving children's attainment through a better quality of family-based support for early learning*. [pdf] Centre for Excellence and Outcomes in Children and Young People's Services. Available at: https://www.basw.co.uk/system/files/resources/basw_111929-1_0.pdf.

Speight, S., Takala, H., Read, H., Robinson, Y. and Fletcher, A., 2022. *The REAL programme Evaluation Report*. [pdf] Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/projects/REAL-Programme-Evaluation-FINAL.pdf?v=1678741153.

Sylva, K., Jelley, F. and Goodall, J., 2018. *Making It REAL: An evaluation of the Oldham Making It REAL project*. [pdf] The Sutton Trust. Available at: https://www.ncb.org.uk/sites/default/files/uploads/files/FINAL%2520DOCUMENT.pdf.

Tracey, L., Bowyer-Crane, C., Bonetti, S., Nielsen, D., D'Apice, K. and Compton, S., 2022. *The impact of the Covid-19 pandemic on children's socio-emotional wellbeing and attainment during the Reception Year*. [pdf] Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/projects/EEF-School-Starters.pdf?v=1678786611.