

Randomised Controlled Trial Evaluation of the White Rose Maths Reception Jigsaw

Evaluation Protocol



Evaluator (institution): National Foundation for Educational Research

Principal investigator(s): Ben Styles (Impact) Helen Poet (IPE)

Template last updated: August 2019

PROJECT TITLE ¹	Randomised Controlled Trial Evaluation of the White Rose Maths Reception Jigsaw
DEVELOPER (INSTITUTION)	White Rose Maths (Trinity MAT)
EVALUATOR (INSTITUTION)	National Foundation for Educational Research (NFER)
PRINCIPAL INVESTIGATOR(S)	Ben Styles (Impact), Helen Poet (IPE)
PROTOCOL AUTHOR(S)	Helen Poet, Ben Styles, Guido Miani, Joana Andrade, Anusha Ghosh
TRIAL DESIGN	Two-arm cluster randomised controlled trial with random allocation at school level
TRIAL TYPE	Efficacy
PUPIL AGE RANGE AND KEY STAGE	Reception pupils (age 4-5)
NUMBER OF SCHOOLS	129
NUMBER OF PUPILS	2580 (20 pupils per school)
PRIMARY OUTCOME MEASURE AND SOURCE	Maths score (PUMA test)
SECONDARY OUTCOME MEASURE AND SOURCE	(1) Early Years Foundation Stage Profile (2) Practitioner confidence in teaching maths

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.2 [<i>latest</i>]		
1.0 [<i>original</i>]	18/08/2020	N/A

¹ Make sure that the project title here matches the title of the document. Please ensure that there is an identification as a randomised trial in the title as per CONSORT requirements.

Table of contents

Study rationale and background	2
Intervention	4
Impact evaluation	12
Implementation and process evaluation	20
Cost evaluation	25
Ethics and registration	27
Data protection	27
Personnel	29
Risks	30
Timeline	31
Appendix 1: Evidence base for Reception Jigsaw modules	33

Study rationale and background

Research suggests that high quality early numeracy education in the early years can have long lasting effects which may help to narrow the gap in achievement throughout life. The EEF Early Years Teaching and Learning Toolkit strand, early numeracy approaches, highlights the importance of professional development in supporting early numeracy approaches. Key areas for improvement include supporting practitioners' knowledge of mathematics; knowledge of children's development and development trajectories in mathematics; and understanding of the kinds of activities which support early mathematical learning.

The reception year, as well as being the first year of primary school, is the final year of the Early Years Foundation Stage (EYFS). Teachers are required to complete the EYFS profile (EYFSP) at the end of the academic year when the child turns five by reviewing the evidence gathered for each child against each Early Learning Goal (ELG) and making a judgement about whether a child is meeting the expected level of development, exceeding the level or not yet reaching the level (emerging)².

In 2017, in response to a consultation on primary assessment in England, the government announced³ that the number and descriptors of ELGs would be reviewed. One of the areas that would be looked at was how the EYFS prepares pupils for key stage 1 mathematics. A recent evaluation⁴ of the pilot of the draft ELGs found that while the pilot schools viewed the draft ELGs positively, there was no consensus about whether reception children would be better prepared for key stage 1 as a result. Furthermore, there was some confusion in schools about whether changes to the ELGs constituted a change to the reception curriculum, for example in relation to the level of emphasis given to shape, space and

² For more information see the EYFSP 2020 handbook
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/858652/EYFSP_Handbook_2020v5.pdf

³
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/644871/Primary_assessment_consultation_response.pdf

⁴
[https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Early_Years_Foundation_Stage_Profile_\(EYFSP\)_Reforms.pdf](https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Early_Years_Foundation_Stage_Profile_(EYFSP)_Reforms.pdf)

measures as this had been removed from the ELGs. It is expected that further guidance will accompany the ELGs once finalised. The proposed ELGs were open for consultation until the end of January 2020⁵ and at the time of writing the responses to the consultation were still being analysed⁶. Once finalised the new ELGs are expected to become statutory in 2021 although schools can use them from September 2020 if they choose.

White Rose Maths (WRM) was developed by Trinity Multi Academy Trust, based in West Yorkshire. They offer a range of schemes of work and maths resources for primary schools; they also deliver training packages. Their signature training package is the 'White Rose Primary Jigsaw'; a five-piece, year-long CPD programme delivered in primary schools to all teachers and support staff for Years 1-6. They have delivered this programme to nearly 100 schools so far.

WRM have received large demand for a training package aimed at reception teachers in response to the proposed changes to the reception year described above. Schools have also fed back to WRM that reception teachers often miss out on training opportunities. More broadly, it is recognised that early years teachers in particular often require support to develop their subject knowledge, knowledge of children's development and pedagogical knowledge. As a result of this, WRM have recently developed Reception Jigsaw which has been piloted in a small number of schools. There has been no formal evaluation of the Reception Jigsaw before, but it is founded in strong principles of early years maths teaching and the modules are based on evidence (see Appendix 1 for the full reference list). WRM have drawn on the What Works Clearinghouse 'Teaching Math to Young Children'⁷ review and practice recommendations and NCETM's progressions for early years⁸. They are also working with Dr Sue Gifford to review the materials and also reviewed the content in relation to the EEF guidance report for maths in the early years⁹. Furthermore when the new ELGs are confirmed WRM will review their training to check that it is in line with the new guidance.

This independent evaluation is necessary to assess impact on pupils' maths attainment and practitioners' confidence in teaching maths to reception age children. The impact evaluation comprises a randomised controlled trial, with school-level randomisation as the intervention is delivered to the whole of reception in each school. Due to the introduction of statutory testing at the start of reception, the trial will use a teacher completed baseline assessment based on their observation of each sampled child rather than imposing the burden of another assessment on pupils at the start of the reception year. The follow-up test will use the Progress in Understanding Mathematics Assessment (PUMA) for reception which will be administered by NFER test administrators. The PUMA is produced by Rising Stars (part of Hodder).

The secondary outcome relating to confidence teaching mathematics will be measured using a survey administered at baseline and at the end of the reception year. The instrument will

⁵ https://consult.education.gov.uk/early-years-quality-outcomes/early-years-foundation-stage-reforms/supporting_documents/EYFS%20reforms%20consultation.pdf

⁶ <https://www.gov.uk/government/consultations/early-years-foundation-stage-reforms>

⁷ <https://ies.ed.gov/ncee/wwc/PracticeGuide/18>

⁸ <https://www.ncetm.org.uk/resources/52500>

⁹ <https://educationendowmentfoundation.org.uk/tools/guidance-reports/early-maths/>

be adapted from that used by Chen et al. (2014)¹⁰ and used recently in the Maths Champions¹¹ trial, with minor adaptations to reflect the setting.

The implementation and process evaluation (IPE) will complement the impact evaluation by gathering important information about the delivery of the Reception Jigsaw and factors affecting its impact.

Further detail about the methods to be used are outlined below.

Intervention

Template for Intervention Description and Replication (TIDieR) Checklist

1. Brief name

White Rose Maths Reception Jigsaw

2. Why (rationale/theory)

White Rose's Reception Jigsaw is a professional development training package that aims to improve maths teaching quality and outcomes in Reception classes. Building on their popular CPD programme, the 'White Rose Primary Jigsaw' (which covers Years 1-6), Trinity Multi Academy Trust have developed a training package for Reception. WRM received a large demand for their Primary training package and feedback from schools that reception teachers often miss out on training opportunities. In response to this, WRM developed the Reception Jigsaw, aiming to improve the specialist maths pedagogical skills and knowledge of reception teachers/ teaching assistants. Other KS1 teaching staff are also invited to attend the sessions (particularly Year 1 teachers), to build on the delivery/ teaching in reception. In particular, Year 1 teachers are invited to attend the training because the development of early number sense and early calculation strategies continues into Year 1. Although the content is predominantly aimed at reception teachers, the training covers how the content can be extended into Year 1. It is also useful for the Year 1 teachers to have an understanding of what has been covered in reception, in order to support those children who need extra support to keep up with the rest of the class. Including both reception and Year 1 teachers (and in some schools, Year 2) provides the opportunity for discussion and comparison of how key learning points may be introduced in reception and in Year 1.

The Reception Jigsaw is underpinned by strong principles of early years maths teaching, specifically: learning through play; opportunities to explore and investigate through the classroom provision; starting from and building on children's interests; using real objects in meaningful contexts to introduce the maths and; the role of the adult in supporting and enhancing learning in the early years through a balance of adult led and child initiated activities. Research shows that high quality early numeracy education has the potential to have lasting positive effects that may help to narrow the gap in achievement throughout life. The Early Years Teaching and Learning Toolkit strand highlights the importance of professional development in supporting early numeracy approaches. Key areas for improvement include supporting practitioners' knowledge of mathematics; knowledge of children's development and development trajectories in mathematics; and understanding of

¹⁰ Chen, Jie-Qi., McCray, Jennifer., Adams, Margaret., Leow, Christine. (2014) *A Survey Study of Early Childhood Teachers' Beliefs and Confidence about Teaching Early Math*. Early Childhood Education Journal (2014) 42:367–377

¹¹ <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/maths-champions/>

the kinds of activities that support early mathematical learning. WRM have drawn on the What Works Clearinghouse 'Teaching Math to Young Children' review and practice recommendations, and NCETM's progressions for early years. The evidence base for each of the five modules is shown in Appendix 1.

3. Who (recipients)

Table 1 below shows which members of staff take part in each element of the jigsaw; these people are the direct recipients of the training.

Table 1: Participants of the elements of training

	Face to face twilight training	Gap tasks	Half-day coaching visit
Reception teachers	Essential	Essential	Essential (in 2+ form entry, participation in each coaching visit will be by one of the teachers, not both/all. Different reception teachers may take part in different coaching visits)
Reception TAs	Recommended	Optional	Typically do not attend
Year1 teachers	Essential	Optional	Typically do not attend
Year 1 TAs	Recommended	Optional	Typically do not attend
Year 2 teachers and TAs	Optional	Optional	Typically do not attend
Maths Lead	Essential	Optional	Essential (although could be for just part of the visit)

Pupils in Reception are the in-direct recipients of the training. Subject to the results of the trial and the security rating, a longitudinal follow-up may take place with the same pupils when they complete Year 1.

4. What (materials)

Specific to the schools participating in the intervention group

- Resources provided in training - There are no specific packs or resources provided to take away that can be used in class. However WRM have a bank of resources available that trainers can share with the school, depending on the needs of the school and what comes up in their discussions.
- Gap tasks – at the end of each twilight session the participants are set a 'gap task' related to the training given, to be completed before the next session (also see section on 'what' below). The gap tasks are listed in the same booklet as the journal provided by WRM.
- Journal to record individual reflections and progress – this is a tool for reflection by those participating in the Jigsaw. It is not collected in or reviewed by WRM although

practitioners might use it during the half day visit by the trainer to talk about their progress.

- Videos which will be made available only to participant schools via a log-in¹². All the videos are of teachers modelling the practice covered in a module. Some will be included in the training sessions, and the training may recommend participants watch specific videos, but the videos are not compulsory.

Available to all schools

- Online resources – publically available to all schools on the WRM website – This includes teacher guidance, schemes of learning, interactive white boards. Awareness of these resources may be higher in the intervention schools due to their ongoing participation in the training. However the control group will also have some awareness of the online resources as they are mentioned as part of the recruitment to the trial.

5. What (procedures, activities or processes used)

The training received by the schools:

- **5 X Twilight in-depth training sessions** (each 2 hours), delivered face-to-face at each school. Each school receives their own training (i.e. not grouped with other schools) – for attendees, see ‘who’ (section above). Sessions are spaced across reception year.
- Sessions will be delivered using slides and handouts of the slides will be provided to schools. Slides will be the same across schools, but discussions and focus will be tailored depending on the school. Trainers are able to add examples based on their own experience, but they are not able to add/take away any slides.
- Sessions will be led from the front, with opportunities for discussions about some of the issues currently faced by the school/approaches taken.
- The training sessions include practical activities and games that schools can use in their own classrooms.

- **5 X Gap tasks** - to be completed between sessions by reception teachers – to encourage implementation - and reflection - of the learning in their teaching.
- Gap tasks are based on the content of the twilights. All of the schools are given the same gap task, but how they approach the tasks is fairly open.
- The gap tasks are discussed with the SLE during the half-day coaching visit (see below).

- **5 X Half day coaching visits** from an Early Years SLE (Specialist Leader of Education)
- These sessions are attended by (one of) the reception teachers. The Maths Lead also usually attends at least part of the coaching session. This is to reduce burden on schools and the need for classroom cover. In 2+ form entry schools the coaching sessions may be shared between the reception teachers.

¹² WRM are investigating whether they will be able to track whether participating schools are viewing the videos.

- The aim of these visits is to support reception teachers in developing effective practice in their own setting, and as such are highly tailored to each school. Schools can also provide feedback on the extent to which they have used the training in their classroom.
- The half day coaching visits are made after the gap tasks and are an opportunity to reflect on the gap task with the SLE.
- These sessions are more tailored than twilight sessions – but are still based on the topic covered in the twilight. The school and SLE agree how to use the time – this could involve coaching, lesson observations, reviewing the organisation of the classroom in relation to maths. Different schools may focus the time differently, for example some schools may opt to skill up a reception teacher to become more expert in maths teaching in the early years, while others may use the time to support newer or less confident reception teachers of maths.
- Twilight training sessions are carried out outside teaching time. Half day sessions are during their school times.

The Jigsaw will be delivered to each school from November 2021 to May 2022. The same person (SLE/trainer) will deliver all the sessions in a particular school – this allows for continuity and relationship building.

The five sessions/modules are:

1. **Developing Early Number Sense:** Focuses on counting principles, subitising, composition of number, comparison and number relationships with gap task around building in activities to subitise through daily inputs and adaptations to the classroom provision.
2. **Creating a Mathematical Classroom:** Themed around attitudes to learning, and incorporating everyday maths through classroom routines. The module also covers the role of the adult in supporting learning and planning for adult led activities, with the gap task revolving around planning a sequence of learning to include a balance of adult led and independent play based activities.
3. **Mathematical Talk and Questioning:** Discusses why talk is important and focuses on developing sustained shared thinking. It also focuses on creating opportunities for talk using open ended questions, examples and non-examples. The gap task asks teachers to trial and observe the quality of talk.
4. **Reasoning and Problem Solving in the Early Years:** Focuses on classroom culture, developing reasoning, and problem solving through games and stories. The gap task is about trialling some of the suggested strategies for developing reasoning and problem solving in the classroom.
5. **Exploring pattern and shape:** Discusses the importance of pattern spotting, progression through pattern, patterns through story, odd and even, doubling, and looking at shapes. The gap task focuses on choosing at least two areas of provision and considering how they could be enhanced to include opportunities for exploring pattern.

6. Who (implementers)

The Jigsaw is delivered by expert trainers from White Rose. Teachers in receipt of training then deliver content to pupils over the course of the reception year.

Who are the trainers?

- Trainers are chosen selectively, both based on applications from potential trainers and 'head-hunted'. Some trainers are full-time employees of WRM. All trainers are qualified teachers and SLEs, and are required to have early years' experience. They have at least 5 years' of teaching experience; some have been teaching for 25 years.
- Trainers are not necessarily maths specialists, but most tend to be. Trainers have a mix of expertise.
- The training programme in Essex operates on a host school/franchise model and WRM train the trainers operating in that area too.

Training of trainers

Who are the trainers, and how is training of trainers carried out?

- The core team at WRM runs the training. WRM make the people being trained ('trainers') aware of the high standards expected. The training of the trainers is delivered by one or more of the core WRM team.
- The train the trainer sessions follow the format of those successfully implemented on the Primary Jigsaw. Trainers will have the opportunity to explore each session in depth, discussing the key pedagogy and research which forms the basis of the training. They will have the opportunity to read and discuss the key background reading which supports each session. In addition to training on the session content, the trainers will also have training on presentation and delivery style.
- As part of the training, the trainers then practice delivering each of the twilight sessions in front of the WRM core team, and WRM provide feedback based on the sessions to the individuals.
- WRM started to recruit trainers during the 19/20 school year in line with the original delivery timetable, and plan to recruit additional trainers in January 2021 so that they are in place before the start of the trial. Trainers will receive some training prior to the start of delivery (in summer 2020, and in spring/summer 2021, depending on when recruited). The remaining training - about how to deliver each of the modules - will be staggered through the 2021/22 academic year to ensure that the delivery of the sessions is not too far removed from the training. This is so that the trainers are trained to deliver a module shortly before they run that module with schools.

Quality Assurance of the trainers' delivery

- Training sessions, delivered by every trainer, are observed by WRM twice a year. Based on their observations, WRM collate their feedback on different aspects of the training. The feedback is then analysed by the WRM team to identify and suggest specific areas for improvement for every trainer.
- As new trainers are being recruited by WRM to deliver the Jigsaw as part of the trial, WRM plan for all newly recruited trainers to deliver a 'practice' session before they deliver to schools in the trial. WRM will observe this practice session and provide feedback and guidance to the trainer as needed.
- Feedback from attendees is collected after every session which is analysed – and used to inform both the course content/delivery across all trainers as well as the line management/development of the individual trainers.
- The trainers have fortnightly meetings with their line manager for supervision and (two-way) feedback.

How long does the training of the trainer last?

- There is one day of introductory training at the start, then for each module there is approximately 1.5 days training per module. The training is intensive and includes the research that the modules is based on. Their delivery of the twilights is observed (by the WRM team) and feedback is provided.
- Training is spaced out over the year – trainers are trained just before they are due to deliver the particular module.

7. How (mode of delivery)

The same trainer delivers all training to the same school. Each trainer goes to between 5 to 10 schools. The school has the trainer's email so they can ask queries.

What is the balance between trainers following the content versus tailoring it?

Twilight sessions mostly follows the specified content, half-days allow more flexibility for tailoring. Trainers are free to add their own anecdotes (also see above).

The participants use what they have learned in the training to change/inform their teaching and delivery of maths in the classroom. The trainers can provide support to do this, as part of the coaching, and if there are questions from the school/teachers then the trainers will support the school with additional requests.

As well as delivering the training and coaching, the trainers observe any changes that have occurred through the year. WRM are able to (lightly) monitor schools' participation in the project during each visit, based on how engaged they are with the content, the extent to which the school has completed the gap task and reflection.

8. Where (setting of the intervention)

The target schools are:

- Primary schools in Yorkshire and surrounding areas (up to 75 miles from Halifax), and schools in Essex or those close to the Essex border in Outer London.
- Those that have not previously had more than two sessions of WRM Primary Jigsaw training or any sessions of Reception Jigsaw training¹³
- Schools who are not participating in any other EEF early years trials.
- Schools with stand-alone Reception classes (i.e. not mixed reception/year 1).
- An adjustment was made to the eligibility criteria in January 2020 relating to the DFE early years professional development scheme (which is aimed at nursery teachers). Schools may participate in this trial *and* the EY professional development scheme so long as the reception teachers in the school do not take part in the maths module of the scheme.

9. When and how much (dosage and duration)

Also see 5

¹³ WRM plans to offer the Reception Jigsaw training sessions as either webinars or face-to-face training during the 20/21 academic year. The half day school support visits will not be included. Any school accessing any of these sessions will be excluded from the trial.

Teachers are encouraged to adapt the training to their own classroom, so implementation will look different in different schools.

10. Tailoring (adaptation to the intervention)

Jigsaw is delivered at each school and there are coaching sessions for each school. The Jigsaw is delivered using a standard set of materials that White Rose have developed, but the school level delivery can be tailored to individual schools if there are particular areas they need support with.

The coaching (half-day) visits are highly tailored to each school (see above).

Teachers can tweak how they adapt their learning based on the level the children are at. The training package provides adaptable ideas for new things that can be introduced, no strict lesson plans are provided.

11. How well (planned) Strategies to maximise adherence and fidelity

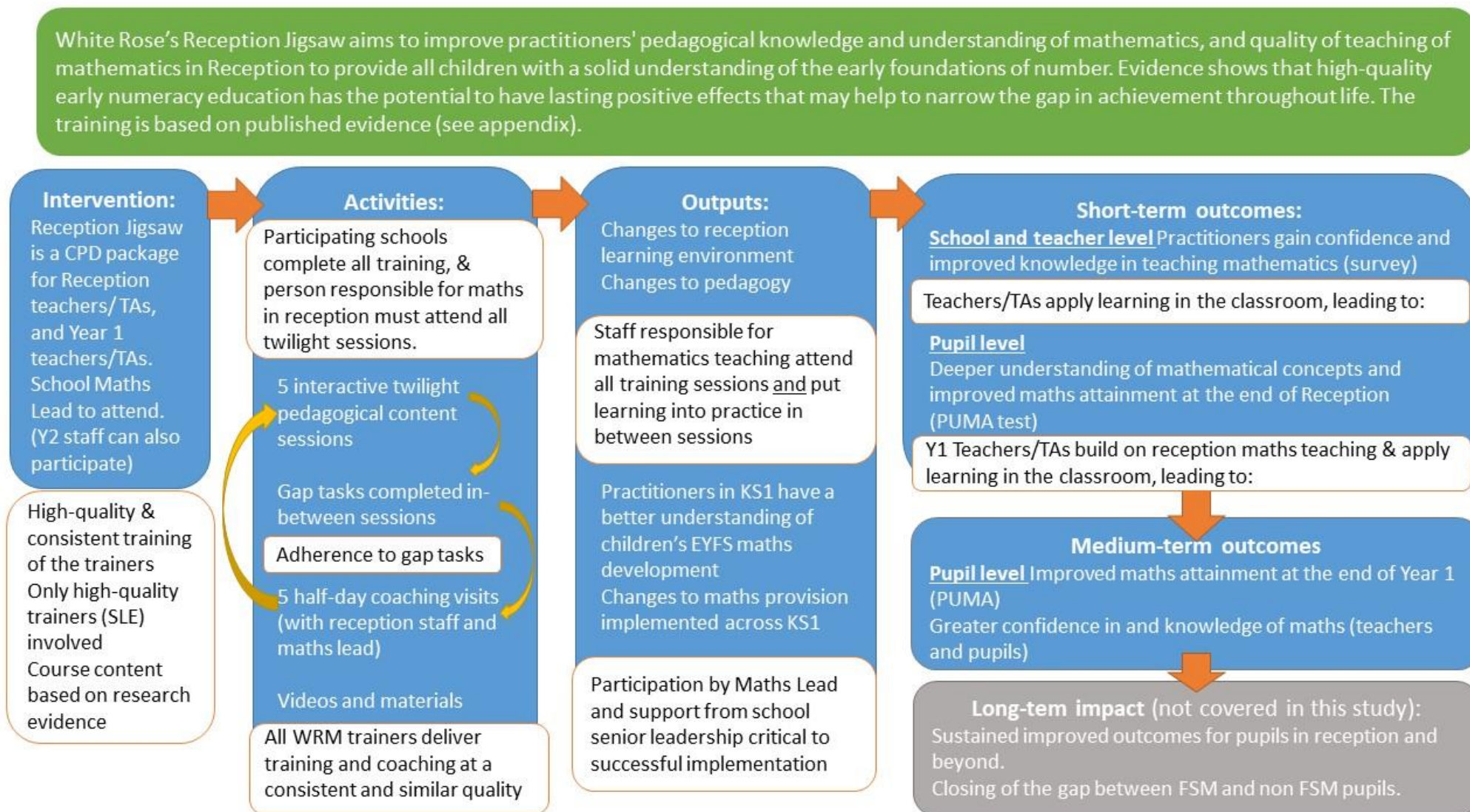
The half day sessions usually allow trainers to gauge engagement and an indication of whether schools are implementing the training. Trainers will feed this information back to the WRM project lead, who will have a follow-up phone call with the key contact at the school to discuss any issues where they arise. WRM project lead will keep a record of any issues and actions taken in response.

WRM collects informal records of where schools are in terms of progress, and whether schools have improved over the year.

LOGIC MODEL FOR THE RECEPTION JIGSAW INTERVENTION

Below is a logic model (Figure 1), which outlines the hypothesised activities, outputs and outcomes. It also summarises the hypothesised underlying causal mechanisms or mediators (in clear boxes). As this is an efficacy trial of an intervention that has not yet been evaluated, the activities, outputs, outcomes and causal mechanisms will be explored through this project, as outlined in this protocol. We do not necessarily know what all of the mechanisms/enabling factors are at this point, nor which are most important. The IPE will aim to identify the most salient factors and we will use this to provide an updated logic model in the final report. The potential mechanisms we are particularly interested in are: the cycle of the training (twilight/gap task/coaching sessions), and the implementation of the learning in the classroom.

Figure 1: Hypothesised logic model for the Reception Jigsaw (mediators shown in clear boxes)



Impact evaluation

Research questions

Primary Question

RQ1: What is the impact of the White Rose Maths Reception Jigsaw intervention on reception children's maths attainment, as measured by PUMA tests at the end of the reception year, compared to 'business as usual'?

Secondary Questions

RQ2: What is the impact of the White Rose Maths Reception Jigsaw intervention on pupils' maths attainment as measured by PUMA tests at the end of Year 1, compared to 'business as usual'?

RQ3: What is the impact of the White Rose Maths Reception Jigsaw intervention on pupils' score on the Early Years Foundation Stage profile (EYFSP), compared to 'business as usual'?

RQ4: What is the impact of the White Rose Maths Reception Jigsaw on practitioners' confidence to teach maths to children in their Reception year, compared to 'business as usual'?

RQ5: Are effects on maths attainment in the reception year (as per RQ1) different for pupils eligible for free school meals (FSM)?

Design

Table 2: Trial design

Trial design, including number of arms		Two-arm, cluster randomised
Unit of randomisation		School
Stratification variables (if applicable)		Geographic area (Yorkshire/Essex)
Primary outcome	variable	Maths attainment
	measure (instrument, scale, source)	End of reception maths score, test total score, PUMA
Secondary outcome (1)	variable(s)	Early Years Foundation Stage Profile (EYFSP)
	measure(s) (instrument, scale, source)	EYFSP, elements to be confirmed in SAP, NPD
Secondary outcome (2)	variable(s)	Practitioners' confidence to teach maths to reception pupils
	measure(s) (instrument, scale, source)	Practitioner confidence, bespoke survey, adapted from Chen et al., 2014
variable		Teacher assessment based on observation

Baseline for primary outcome and secondary outcome (1)	measure (instrument, scale, source)	Observation checklist, 0-20, bespoke instrument
Baseline for secondary outcome (2)	variable measure (instrument, scale, source)	Practitioners' confidence to teach maths to reception pupils at the start of the academic year Practitioner confidence, bespoke survey, adapted from Chen et al., 2014)

Randomisation

Randomisation will be stratified by geographical area using only two strata: Yorkshire and Essex. It will be carried out using R and code will be appended to the report. Test administrators will be blind to group allocation but analysts will not.

Participants

Within each school different members of staff participate in the different elements of the training, as shown in Table 1. As outlined above, schools are eligible for the trial if they meet the following criteria:

- Primary schools situated within a 75 mile radius of Halifax, and schools in Essex or those close to the Essex border in Outer London.
- Those that have not previously had more than two sessions of WRM Primary Jigsaw training or any sessions of Reception Jigsaw training¹⁴
- Schools who are not participating in any other EEF early years trials in 2021/22
- Schools with stand-alone Reception classes (i.e. not mixed reception/year 1).

Schools may participate in this trial *and* the DFE EY professional development scheme so long as the reception teachers in the school do not take part in the maths module of the scheme.

WRM will recruit schools to the trial, and will run a number of recruitment events to raise awareness of the project and answer questions from interested schools. WRM have capacity to deliver the intervention to 35 schools in Yorkshire and 20 in Essex. Therefore, they will aim to recruit 85 schools from Yorkshire and surrounding areas (35 intervention, 50 control) and 50 from Essex and surrounding areas (20 intervention, 30 control).

To minimise the burden on schools, NFER will randomly select 20 pupils from each school to take part in the testing (although all will receive the intervention, as it is delivered to the whole class). This random selection of pupils will be stratified by whether or not they are eligible for FSM to ensure adequate representation of this subgroup within each school. The pupil sampling will take place in August 2021, once we have received pupil lists from schools and before the baseline measurement (observation checklist in the Autumn term).

Recruitment documents will consist of the following:

- Memorandum of Understanding (prepared jointly)

¹⁴ WRM plans to offer the Reception Jigsaw training sessions as either webinars or face-to-face training during the 20/21 academic year. The half day school support visits will not be included. Any school accessing any of these sessions will be excluded from the trial.

- Privacy Notice for parents (prepared by NFER)
- Privacy Notice for schools/teachers (prepared by NFER)
- School information letter (prepared jointly)
- Parent information leaflet (prepared jointly)
- Parental letter to opt out of data sharing (prepared jointly)

WRM will begin contacting schools in November 2020¹⁵. When a school indicates their interest to WRM, they will send them the school information sheet, including links to the privacy notices, and invite them to a recruitment event. After the event they will share with them the Memorandum of Understanding for the school to sign. When the school sends a signed Memorandum of Understanding back to WRM, WRM will record them as recruited to the trial.

At the beginning of May 2021 WRM will share the list of recruited schools with NFER. NFER will then contact the schools in late May and send them, via a secure school portal, the Parent information leaflet, Parent withdrawal from data processing letter, and the Parent privacy notice. The school will be advised to print the documents out for all pupils who will participate in the trial, to be sent home to parents. We will ask schools to give the withdrawal form to parents in May/early June, before we start collecting pupil data, so that we can withdraw pupils before we collect pupil data. This is also to avoid selecting pupils for baseline that will then withdraw.

If a parent withdraws their child from data processing at any time, the school will be required to inform NFER, who will then delete the pupil's data. At the beginning of June NFER will share a data collection template, requesting that the schools return all pupil and practitioner data within three weeks.

Families that join the school at a later stage will still receive all the documentation, and will be able to withdraw their pupils at any point.

NFER will collect the following pupil data:

- names
- date of birth
- Unique Pupil Number (UPN)
- Free School Meal eligibility (FSM) (used to sample)

NFER will collect the following practitioner data:

- names
- contact details
- job role

Once the data has been received, NFER will prepopulate the baseline observational checklists and prepare the baseline practitioner survey for administration in September 2021. The observational checklist template will be sent to schools via the secure online portal, while each practitioner will receive a link to the survey, to be completed online. (For

¹⁵ As noted in Table 8 (timeline) this project was delayed by one academic year as a result of the school closures due to the Covid-19 outbreak in 2020. This was after WRM had recruited over half of the schools for a Sept 2020 delivery, however the risks to delivery in 20/21 were felt to be too great and a delay was agreed. Consequently the recruitment period is planned to recommence in November 2020 – the schools that had already agreed to take part will be eligible to take part if they would like to. These schools will be contacted first when recruitment recommences

further details about these instruments, see the outcome measures section below). Schools will be given a six week window in September/October 2021 to complete both instruments (and this can only be done after the school has completed the RBA). Schools will need to upload their observational checklist via the secure online portal once completed. Once the completed instruments are received from the school, schools are put forward to randomisation. Randomisation will take place in mid-October, in order to notify schools of group allocation in late October, so that delivery can start soon after the October 2021 half term.

Sample size calculations

Table 3: Sample size calculations

		OVERALL	FSM
Minimum Detectable Effect Size (MDES)		0.22	0.27
Pre-test/ post-test correlations	level 1 (pupil)	0.40	0.40
	level 2 (class)		
	level 3 (school)		
Intraclass correlations (ICCs)	level 2 (school)	0.17	0.17
	level 3	n/a	n/a
Alpha¹⁶		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		two-sided	two-sided
Average cluster size		20	5
Number of schools¹⁷	Intervention	55	55
	Control	74	74
	Total	129	129
Number of pupils	Intervention	1100	275
	Control	1480	370
	Total	2580	645

As with all EEF trials, it may make sense to establish a sampling frame of high FSM schools to increase the power of FSM sub-group analyses. These calculations assume 25% of pupils are eligible for FSM i.e. considerably higher than the national proportion.

We have used parameters from the Maths Champions evaluation (Robinson-Smith *et al.*, 2018)¹⁸ which used CEM's ASPECTS assessment. This had an ICC of 0.17. Using Early

¹⁶ Please adjust as necessary for trials with multiple primary outcomes, 3-arm trials, etc., when a Bonferroni correction is used to account for family-wise errors.

¹⁷ Please adjust as necessary, e.g., for trials that are randomised at the class level.

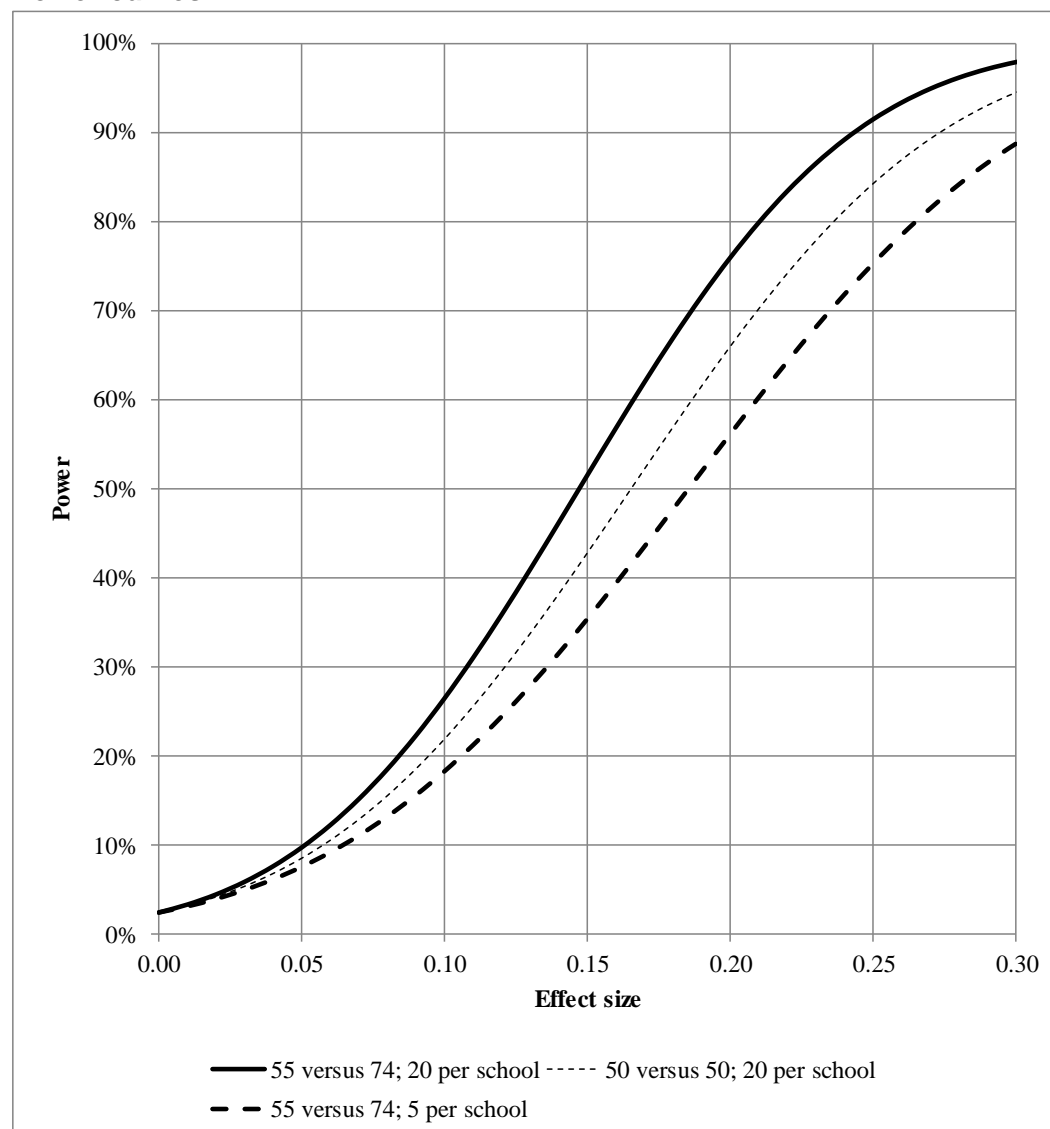
¹⁸ <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/maths-champions/>

Years Outcomes at baseline will result in a lower correlation than a properly developed test so we have assumed a correlation of 0.4.

It transpires that considerable cost savings can be made by only testing a random sample of 20 pupils per school since this is the number of small-group tests that an NFER test administrator can accomplish in a day's visit. Using these assumptions and after discussions with the developer concerning their capacity to deliver to more than 50 schools, a sample size of 55 intervention schools and 74 control schools was agreed. This design has 80% power to detect an effect size of 0.22. These are achieved sample sizes i.e. schools with both baseline and follow-up data. The trial should recruit around 5% extra schools, say 135, to allow for dropout at baseline before randomisation. This way, we should randomise at least 129 schools, always 55 to intervention and the remainder to control. The extra schools will allow for any that refuse testing at follow-up, which should be few in number due to the use of test administrators.

Power calculations were completed using a bespoke Excel spreadsheet.

Power curves



Outcome measures

The design is a two-arm trial with school-level randomisation. The primary outcome is of maths attainment and will be measured on Reception children, with a possible follow up at

the end of Year 1. This relates to the short- and medium-term pupil-level outcomes specified in the logic model above.

Secondary outcomes of (1) the Reception pupils' Early Years foundation Stage Profile (EYFSP) and (2) practitioners' confidence in teaching maths (also relating to the short-term outcomes in the logic model) are outlined below.

Baseline measures

The first live national administration of the NFER's new [Reception Baseline Assessment](#) (RBA) will occur for [Early Adopter schools](#) in autumn 2020. It is intended for administration within the first six weeks of a child starting school, whenever that might be during the year¹⁹. At this stage (June 2020), we understand that the RBA data will not be available to researchers via the NPD²⁰.

Sample size calculations for this trial imply that a baseline measure is required. Realistically, the timing of the assessment prevents us from administering a separate baseline with the pupils for this trial due to the burden on schools and their new pupils. Therefore, we will devise an observational checklist for reception teachers to complete. This checklist will be based on the Early Years Outcomes²¹ and the Early Learning Goals²², and aligned to information that reception teachers are likely to be gathering and recording anyway. We will ask schools to complete the checklist for each child once they have completed the RBA²³. As this project has been moved back a year to 21/22, we will be able to review our planned process and timetable against the administration of the first year of the RBA (in 20/21).

Randomisation will take place in mid to late October, after the RBA has been completed and after schools have completed the observational checklist. Pupil data collection will also occur in advance of randomisation. This will allow intervention delivery over nearly a whole academic year which WRM are amenable to. We recognise that this may cause some difficulty for schools who cannot plan for the academic year over the summer as they do not know their allocation. We propose to work closely with WRM to identify how we can support schools through this period, providing schools at recruitment/MOU stage with detailed information about staff and resource requirements so that schools can develop a plan ready to put into action if they are allocated to intervention. WRM will request availability dates for the first training session from all schools, ahead of randomisation. This will facilitate a prompt start of the delivery, as dates will only need to be confirmed by schools allocated to the intervention group. We will also stratify the randomisation geographically to avoid clumping and to aid intervention delivery.

Primary outcome

We considered test choice for follow-up closely, with three main criteria driving selection:

¹⁹ Although the Early Adopter year will be run after October 2020 half term due to the 2020 school closures, the plan for 'normal' school years is that the vast majority of children would experience the RBA before October half term. The project timetable is designed assuming the RBA will be administered before October half term in 2021. This will be reviewed.

²⁰ Based on the information from the privacy notices for the RBA, available here:

<https://www.nfer.ac.uk/for-schools/participate-in-research/information-about-the-reception-baseline-assessment/>

²¹ https://www.foundationyears.org.uk/wp-content/uploads/2012/03/Early_Years_Outcomes.pdf

²² https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/790580/EYFSP_Handbook_2019.pdf

²³ Not immediately afterwards, due to requirements of the administration of the RBA.

- Recent UK standardisation
- Content alignment with current national curriculum and EYFS
- Practicality of administration.

Many of the tests identified on the EEF Early Years Measures database are not standardised for the UK and many rely on one-to-one assessment. Though clearly a reliable method of assessment for the Early Years, one-to-one is difficult for schools and costly to administer. We considered three assessment measures: Hodder's MALT, Hodder's Progress in Understanding Maths (PUMA) and GL Assessment's Progress Test in Maths (PTM). We discounted PTM because it does not start until age 5. Our suggestion for the primary outcome measure is the PUMA, which has a test version designed and standardised for children of 4 years 10 months (i.e. designed to be taken after half-term in the summer term of reception). This test was written and standardised more recently than its close competitors and is targeted to the current primary curriculum; it also links to the Development Matters strands in the EYFS and ELG ranges. It received one star for its psychometric properties and two stars for ease of implementation on the EEF Early Years Measures database²⁴.

PUMA will be administered in June/July 2022. The assessment will be administered by experienced NFER Test Administrators²⁵ who will be recruited on the basis of their Early Years experience and trained specifically for the small-group (5-6 pupils) testing regime required. Administrators and markers will be blind to group allocation.

Secondary outcomes

EYFSP

The first secondary outcome will be analysis of the EYFSP maths results, using data from the NPD linked via UPN obtained at the outset from schools. The EYFSP is recorded at the end of the reception year, around the same time we will be administering the PUMA tests. The benefit of the additional EYFSP analysis is that it will be aligned with the Early Learning Goals and includes some elements that the PUMA does not (such as manipulatives, which are a feature of the intervention). However, it remains a secondary outcome since it is teacher assessed and not of high reliability.

The EYFSP is a measure of a child's attainment in relation to 17 early learning goals²⁶ prior to the age of five. The latest version of the scale has only two maths measures (G11 and G12). Due to concerns that two measures with a scale of three points each may not be a sufficiently informative measure when summed, we will conduct a reliability analysis to decide whether to proceed with its use as an outcome. If its Cronbach's alpha is below 0.7, we will not proceed with this analysis.

Practitioners' confidence in teaching maths

The other secondary outcome will be the impact of the intervention on practitioners' confidence to teach maths in the reception year. We will use a pre- and post-survey of teachers, teaching assistants and the school maths coordinator. The survey will be based on the survey developed by Chen et al. (2014) and used recently in the Maths Champions

²⁴ <https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluating-projects/early-years-measure-database/early-years-measures-database/>

²⁵ We will review whether sending Test Administrators into schools is appropriate and within government guidance on social distancing nearer the time.

²⁶ https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/790580/EYFSP_Handbook_2019.pdf

trial²⁷, with minor adaptations to reflect the setting (mainly related to terminology, i.e. to say reception/school rather than nursery). The survey will be administered at the same time as the pre-and post-primary outcome activities.

Compliance

Compliance for the intervention will be defined in terms of the number of training sessions completed by each school and feedback from the trainers about schools' engagement with the programme.

A school will be considered to have completed the intervention if it participates in nine or more sessions out of the ten available (five Twilight in-depth training sessions and five half-day coaching sessions); this binary measure will form an optimal compliance indicator in the CACE analysis (see below). In addition we will run a separate CACE analysis using a continuous measure of compliance (0-10 sessions, as above). While individual teacher attendance at the twilight sessions should be high, the nature of the coaching sessions is such that we would not expect all individuals to attend all coaching sessions (also see Table 1). (WRM will collect attendance registers with details of each participant, including their name and role in the school at all training sessions. We will draw on this information for the case studies in the IPE.)

While adherence to gap tasks will not strictly be used as a measure of programme compliance (i.e. we will not be collecting the gap tasks in), the trainer's perception of the level of school engagement will provide information about the extent to which the schools are applying the learning in the classrooms, and to which the intervention is reaching the pupils. This information will be based on the extent to which they have completed the gap tasks and are observed to have implemented the learning. It will be collected in a reflection sheet to be filled out by the trainer at the end of each half-day session (see Tool 3 in IPE section).

Analysis

The main analysis for this trial will be intention-to-treat, and will follow the EEF analysis guidance²⁸. It will compare PUMA in the intervention arm with control. We will use the total score from the observational checklist completed at baseline as a covariate, along with any stratifiers, in a multilevel model that takes into account the cluster randomisation. Hedges *g* will be calculated using total variance from a model without covariates. 95% confidence intervals will be computed for each effect size. Analysis will be pre-specified in a detailed Statistical Analysis Plan.

Child-level attrition is unlikely to be a significant issue in this trial due to the use of test administrators. School-level measurement attrition may be a concern, due to the new testing burden in this age group, which will be added to by participation in this trial. Using NFER's regular keeping in-touch strategy, employing NFER test administrators, providing detailed pupil-level results feedback and using a uniform incentive for both intervention and control testing will help to minimise attrition. Despite this, it will be pertinent to over-recruit slightly to allow for withdrawal both before and after randomisation. Should attrition be an issue at analysis stage, we will adhere to the EEF guidance on missing data analysis.

²⁷ See footnotes 10 and 11, above

²⁸

https://educationendowmentfoundation.org.uk/public/files/Evaluation/Writing_a_Protocol_or_SAP/EEF_statistical_analysis_guidance_2018.pdf

EVERFSM6 from the NPD will be used as the identifier for the main subgroup analysis.

Two sets of Complier Average Causal Effect (CACE) analyses will be carried out using fidelity measures (binary and continuous) developed from the school-level participation data and the school engagement ratings by trainers (see compliance and IPE sections).

Longitudinal follow-ups

When the reception pupils are in Year 2, the KS1 statutory tests will no longer be mandatory, and therefore the next time-point for administrative data would be the end of Year 6 (2028). A more timely follow-up would be to re-administer PUMA at the end of Year 1 (in 2023). This timing would have the added benefit of including the additional influence of Year 1 teachers who were also trained alongside Reception teachers. This follow-up will only proceed subject to the EEF guidance for longitudinal analysis²⁹. For example, if the intervention was implemented with low fidelity or there was evidence of experimental effects, longitudinal analysis may not be appropriate.

These results would be analysed using a separate multi-level model as, although it is the same test as for the primary outcome, it will be a different version (Y1 as opposed to Reception).

Implementation and process evaluation

Research questions

The implementation and process evaluation (IPE) will complement the impact evaluation. It will gather important information about the delivery of the Reception Jigsaw and factors affecting its impact. It will cover the complete set of eight dimensions and five implementation factors described in the IPE introductory handbook³⁰, with a special focus on fidelity/adherence, Participation/implementation responsiveness and adaptation, quality, dosage, implementation support factors, and programme differentiation.

The IPE will seek to answer the following questions:

IPE_RQ1: Was the Reception Jigsaw delivered as intended in terms of dosage, nature, and quality?

IPE_RQ2: How well did the participants (teachers/support staff (reception and other KS1), maths coordinator, senior leaders, and then also pupils) engage with the Jigsaw? Were there any implementation challenges faced? If so what were they and to what extent were they overcome?

IPE_RQ3: Was the quality of training, support, and intervention materials provided by the developer adequate? Was preparedness and confidence of staff delivering the intervention at the right level? If not, why?

IPE_RQ4: To what extent do participants feel the intended outcomes of the programme are being achieved for children, practitioners, and the school? How are they being achieved?

²⁹

https://educationendowmentfoundation.org.uk/public/files/Grantee_guide_and_EEF_policies/Evaluation/Writing_a_Protocol_or_SAP/longitudinal_guidance.pdf

³⁰

https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Handbook.pdf

IPE_RQ5: What does business as usual (BaU) consist of for the comparison group?

Research methods

The IPE will involve a range of research methods described below.

IDEA workshop – IPE_RQ1, IPE_RQ3, IPE_RQ5

The IDEA workshop was held on 13th November 2018, and provided an opportunity for the White Rose Maths development team and the NFER evaluation team to explore the intervention in depth, and to develop an effective implementation and process evaluation (IPE) plan. Plans for the instruments, interview schedules and observation tools were confirmed. Attendees also developed the Template for Intervention Description and Replication (TIDieR) framework, discussing key features of the intervention; then explored the Theory of Change, reflecting on the developer team's original ToC and ensuring a mutual understanding of the research process and its aims.

In addition, the fidelity/compliance criteria were discussed, including identifying the critical components of implementation and how to identify 'success' within each. This included monitoring requirements for both WRM and NFER. The meeting also provided an opportunity to understand how the coaching sessions may vary and how this can best be captured for the purposes of fidelity measurement in the implementation diary.

Tool 1: Practitioner Surveys – IPE_RQ1, IPE_RQ2, IPE_RQ3, IPE_RQ4, IPE_RQ5

Baseline Proforma

A short proforma will be sent to schools after they have returned the MOU. This will ask schools to indicate what their Business as Usual is in terms of numeracy pedagogy in Reception. This will allow us an understanding of usual practice across all schools.

End-point Proforma

A second proforma will be sent to control schools at the end of the 21/22 academic year. This proforma will ask about the maths CPD that took place for reception/KS1 staff during the trial period.

Endpoint Practitioner Survey

A concise set of survey questions about participation in the intervention will be designed. This will be administered online to intervention schools, appended to the end-point survey for the secondary outcome on confidence teaching maths (summer 2022), for reception teachers and teaching assistants, and maths leads. A separate, short version of the process questions will be sent to Year 1 and Year 2 practitioners, and headteachers (as they will not receive the survey for the secondary outcome). As the intervention targets change at a whole-school level the survey will consist of slightly different versions targeted to Reception, Year 1 and Year 2 teachers and Teaching Assistants (TAs), mathematics coordinators and head teachers. The survey will cover the following areas:

- **Dosage:** The extent to which the training was applied when teaching
- **Quality:** Training and support, preparedness to deliver; school implementation support factors

- **Reach:** Which pupils receive the learning from the intervention; do any pupils receive more or less? (e.g. do they target the learning at particular pupils or apply in whole-class approaches?)
- **Responsiveness:** Do teachers/TAs like delivering the intervention? Perception of change in attitudes, behaviours and practice at class and school level. How do pupils respond to the intervention?
- **Programme differentiation:** How distinct is the programme from existing practice?
- **Adaptation:** What changes are made to the intervention during implementation? What adaptations are made during delivery by teachers, and why? How much and to what extent does the programme differ between schools?

We will design all surveys to be as light touch as possible to reduce the burden on schools.

Tool 2: Training Observation and Registers – IPE_RQ1, IPE_RQ2, IPE_RQ3

The training of the newly recruited White Rose Maths early years specialists will be observed in person to enable an understanding of the trainer's role as a coach and mentor to the school. We will then observe³¹ one CPD (twilight) session delivered by four of the trainers as part of the case study visits. This is in order to gain a measure of any differences in quality, engagement, style of delivery or adaptation of content at the training sessions. For each of these observations we will create bespoke observation schedules.

For both the twilight sessions and half-day coaching sessions, a register of attendance will be collected by WRM. This will be used to establish an indicator of dosage/reach.

Tool 3: Trainer reflection sheet – IPE_RQ1, IPE_RQ2, IPE_RQ4

As part of the intervention, participants are asked to complete gap tasks by completing a booklet. This booklet also has space for reflection. WRM do not collect the booklet back in; it is a tool to support the learning and reflection of participants – this would continue to be a part of the intervention in any future roll-out of the jigsaw. As such, we do not believe it is appropriate to impose collection of these booklets as part of the trial, not least because knowing that the booklet will be collected and reviewed is likely to affect how participants complete and use the booklet. (If they wish, teachers may wish to draw on their recordings in their booklets when answering surveys and interview questions.)

However, we do need to understand the sort of support schools are receiving, and how and when the coaching is being implemented. To do this we will design a short form for the trainers to complete at the end of each of the half-day sessions. These forms will capture the trainers' perceived engagement of schools including completion of gap tasks by schools and implementation of the learning in classrooms. We will also ask the trainers to indicate the type of support they provided during each coaching session (e.g. revisiting content from the twilight sessions, more in-depth support towards developing in-school experts and/or support for implementation in the school). Questions will primarily be close-ended such as in the form of Likert scales, to reduce burden. This data will be collated by WRM with the information about school/participant attendance at the twilight sessions and half day coaching sessions.

³¹ We will review the suitability of observations in line with the current Covid-19/social distancing guidance at the time the observations are due to take place (spring/summer 2021).

The trainers' perceptions of schools' engagement will allow us to capture the extent to which schools participate in and implement learning from the intervention programme and will contribute to the compliance measure described above.

Tool 4: Case Studies and Interviews – IPE_RQ1, IPE_RQ2, IPE_RQ3, IPE_RQ4, IPE_RQ5

We will conduct a set of case studies³² operating longitudinally over the course of the intervention year. This will be carried out in four schools, at three time points (participants at each point are outlined below). The rationale behind the case studies is twofold. First, we anticipate a fair amount of turbulence in early years due to changes to statutory testing, which is likely to have an impact on both confidence and the amount of innovation in schools during this period, which could well have an impact on engagement with this intervention. Second, this intervention could have an impact at whole-school level, meaning there are various inter-dependencies that could have an effect on impact that would otherwise be difficult to capture (e.g. strength and style of leadership, school policy, amount and type of TA support, level of maths coordinator support and engagement).

The focus of the case studies will therefore be around dimensions of fidelity, dosage, quality, reach, responsiveness, programme differentiation, adaptation, and costs. The case studies will also enable a deeper understanding of how the coaching visits have worked for the school, and how closely aligned they are with the school's perceived needs. The case-study design is such that we will observe both twilight sessions and the half-day coaching.

The case-study design is as follows:

- Four schools selected at random, with geographic spread
- Checkpoint 1 (Autumn 2021, at the time of the first twilight session)
 - Semi-structured one-to-one interviews of approximately 30 minutes in length³³ with reception, Year 1 and Year 2 teacher (as appropriate, depending on intended participation in that school), and maths lead, in person
 - These interviews will explore the reasons why the school signed up for the Jigsaw, including the school's previous approach to EY maths provision and intentions for implementation. At an individual and school level we will explore confidence to teach maths, attitudes towards the training, and current classroom practice.
 - Observation of the first twilight session
- Checkpoint 2 (January/February 2022)
 - Interviews with staff as CP1, on telephone
 - Interviews at checkpoint 2 will cover perceptions of and attitudes towards the training and support provided (including perceptions of how pupils have responded), adaptation and implementation of learning, and time spent.
- Checkpoint 3 (Summer 2022)
 - Interviews with staff as CP1 plus head teacher, in person
 - Observation of the last half day coaching session.
 - Interviews in the final checkpoint will reflect on the provision over the year, changes in confidence and practice (individually and across reception and KS1), changes in the classroom including pupil response, perceived comparison to any other maths training experienced previously, time/resource spent (re: cost evaluation).

³² The case studies are intended as face to face visits to the schools. In light of the Covid-19 guidelines, we will review the suitability of face-to-face visits nearer the time, as the guidance is updated.

³³ The interviews with reception teachers and the Maths Lead are likely to be longer, and interviews with other KS1 staff are likely to be shorter, reflecting their relative involvement in the programme.

In addition to the longitudinal case studies, we will also conduct two best-practice case studies. These will be carried out in two schools at a single time point towards the end of the intervention window (spring/summer 2022). These case studies will include interviews with the Year 1 and Year 2 teacher, as well as the maths lead and head teacher. The purpose of these studies is to explore how a small selection of schools identified by the developer as an example of best practice has implemented the training. This will enable us to understand an ideal environment for the intervention, whether these settings have still encountered any common barriers and, if so, how they have overcome them.

We will also conduct two to three interviews with key members of the delivery team at the end of the intervention year. This will enable us to establish whether there have been any changes to the intervention during the trial due to the requirements of the evaluation, or for any other reason.

Analysis

Table 4 outlines how the IPE will be analysed and the research question(s) each element seeks to inform. The methods have been selected to answer the research questions drawing on the experiences and perspectives of different stakeholders within schools. Information about the type and extent of maths CPD in both the intervention and control groups will provide evidence about the extent to which the key enabling factors hypothesised in the logic model hold true.

The survey will be analysed using descriptive statistics (frequencies and select cross tabs which will be pre-specified).

The qualitative data will be analysed initially using a top-level coding frame(s), developed from the semi-structured interview schedule. The interview schedules have not yet been developed but are intended to cover the topics outlined in the section above. After the first round of coding, sub-codes will be created and assigned to the text, as the data is analysed and themes emerge. We will analyse the data across roles within a school, as well as across the case study units. The small sample size of the qualitative work means that analysis by school type (or other key characteristics) is not appropriate in this project.

Table 4: IPE methods overview

Research methods	Data collection methods	Participants/ data sources (type, number)	Data analysis methods	IPE Research questions addressed	Implementation/ logic model relevance
IDEA workshop	TIDIER framework; logic model completion	WRM team NFER team	Descriptive analysis	RQ1, RQ2, RQ3, RQ5	Context
Business as usual	Baseline Proforma (online)	All schools (baseline) Control schools (end-point)	Descriptive statistics; correlations	RQ5	Business as usual
Practitioner Surveys	Endpoint Practitioner Survey (online)	Intervention schools	Descriptive statistics; correlations	RQ1, RQ2, RQ3, RQ4, RQ5	Fidelity, Dosage, Quality, Reach, Responsiveness, Programme differentiation,

					Monitoring control/ comparison, adaptation, costs
Training Observation and Registers	Structured observations	WRM team NFER team/ 5 observations as part of longitudinal case studies	Deductive coding; within-case analysis; cross-case analysis	RQ1, RQ2, RQ3	Fidelity, Dosage, Quality, Reach, Responsiveness, Adaptation, costs
	Attendance register	CPD attendance registers	Frequency counts	RQ1, RQ2	Fidelity
Trainer reflection sheet	Trainer reflection sheets	Trainers	Descriptive analysis	RQ1, RQ2, RQ4	Fidelity, Dosage, Reach, Adaptation
Case Studies and Interviews (longitudinal)	Semi-structured one-to-one interviews	4 schools, with interviews at 3 time points (Autumn 2021, Jan/Feb 2022, Summer 2022) with Year 1 and Year 2 teacher, maths coordinator, and 1 time point with the head teacher	Inductive /deductive coding; thematic analysis; within-case analysis/ triangulation; cross-case analysis	RQ1, RQ2, RQ3, RQ4, RQ5	Fidelity, Dosage, Quality, Reach, Responsiveness, Programme differentiation, adaptation, costs
	Observations Telephone interviews				
Case Studies and Interviews (best practice)	Semi-structured one-to-one interviews	2 schools, 4 staff, 1 time point	Inductive coding; thematic analysis; within-case analysis/ triangulation; cross-case analysis	RQ1, RQ2, RQ3, RQ4, RQ5	Fidelity, Dosage, Quality, Reach, Responsiveness, Programme differentiation, adaptation, costs

Cost evaluation

We plan to collect information on the pre-requisite, set-up and ongoing costs to schools of being involved in the reception jigsaw.

For this trial, we think three areas will be particularly important to explore:

- i) the set-up and ongoing ‘administrative/programme-level’ costs to a school of being involved with Jigsaw (e.g. costs of training)
- ii) the actual and relative costs of participating in the Jigsaw
- iii) business as usual costs (usual *and* actual BaU costs to schools for delivering maths teaching in reception and maths CPD during the trial period).

For the purposes of this trial, the Reception Jigsaw is being run as a complete programme of all five modules, and this is what the cost evaluation will be based on as per the EEF cost guidance.

Rather than gathering costs information/data from all schools on all of the above, we envisage collecting data from a sample of schools or from data collected by WRM to get

information where WRM feel that most schools do the same/little variation. For items where we anticipate larger variation we will collect information from all schools. For example:

- The number and length of training sessions will be similar across all schools, although the number of staff attending the twilight training is likely to vary. We will be able to collect this information from the WRM attendance registers.
- The implementation (extent and depth) of learning may vary by school and so we envisage collecting information about the time taken to implement the learning within the classroom(s) (beyond the time spent with the WRM trainer/completing the gap tasks). The implementation costs would therefore be collected via the reception teacher survey, with further explanatory detail provided through the case studies and jigsaw manager interviews.
- Usual maths CPD spend will vary by school. We will capture this using the baseline pro-forma that all schools will be asked to complete, which will ask about schools' usual maths CPD. At the end of the intervention we will collect actual BaU activity during the trial period from the control schools, including top-level information on about actual maths CPD for reception/KS1 during the trial.
- Our assumptions for this trial are to calculate the costs of Jigsaw against the costs of no Jigsaw (just usual maths teaching). We will include some description of control schools' actual spend on maths CPD to illustrate the range of schools' spend on maths CPD.

Table 5 lists the resources to be explored in the cost evaluation, the data sources, and whether we will collect the information from all schools or a sample.

Table 5: Cost evaluation list of potential resources and evaluation sources

Category	Item	Evaluation data source(s)	Scope - collect from all, or a sub-sample etc
Personnel for training	Attendees at the five twilight sessions (see table 1 of protocol) For teachers this would be within their directed time (training and meetings included in this), not extra. TAs don't normally get paid for training and meetings so there may be an extra cost there.	WRM attendance registers, sessions fixed at 2 hours each Case studies: to need to check how/if TAs are reimbursed (TOIL or financially)	All intervention schools via WRM More detail from a sample of case study schools
	Reception teacher participation in half day (approx. 3 hours) coaching session x5 / teacher cover for participation	WRM attendance registers Case studies: to check how schools cover the coaching time (internally or paid supply)	All intervention schools via WRM. More detail from a sample of case study schools
	Completion of gap task This involves trying something out from training and should be part of the normal planning of lessons, so teachers would be planning this instead of planning other maths practice that would have happened anyway.	WRM trainer reflection sheet (completed gap task/not completed) also collect from case studies re length of time to complete gap tasks compared to usual practice	All intervention schools via WRM More detail from a sample of case study schools

Personnel for preparation and delivery (implementation)	Reception teachers (/TAs?) prepare and deliver maths sessions. This would be instead of the maths sessions they would have run anyway.	Case studies to explore how this compares to usual practice	Case study sample schools
	Reception teachers (/TAs?) prepare and set up maths areas This would be instead of the maths areas they would have run anyway.	Case studies to explore how this compares to usual practice	Case study sample schools
Training and programme-level costs	Reception Jigsaw fee Plus as estimate of expenses (mileage) as this is charged to the school in usual delivery (examples for short/long distance from hub/delivery school)	From developer (WRM) - market price, not trial price	Developer
Facilities, equipment and materials	<i>[no resources or equipment required. WRM provide some online materials. Schools may choose to buy resources, but WRM work with what the school has got & don't recommend or require particular resources - more about the skills.]</i>	<i>[explore whether anything additional was purchased – case studies only]</i>	Case study sample schools
BaU costs	Usual costs of maths CPD for reception/KS1	Baseline school pro-forma	All schools
BaU costs	Actual costs of Maths CPD in control schools during trial period	End-point control school pro-forma	All control schools

Ethics and registration

The trial will be designed, conducted and reported to CONSORT standards (<http://www.consort-statement.org/consort.statement/>) and registered on <http://www.controlled-trials.com/>. The evaluation will be conducted in accordance with NFER's Code of Practice, available at: <http://www.nfer.ac.uk/nfer/about-nfer/code-of-practice/nfercop.pdf>. NFER, White Rose Maths and EEF will work together to ensure each organisations' policies can be applied in practice.

Ethical agreement

Ethical agreement for participation within the trial will be provided by the headteacher of the school. Parents will be provided with full details about the intervention, and will be given the opportunity to withdraw their child from testing and data processing if they have objections to this.

All data gathered during the trial will be held in accordance with the data protection framework established by the Data Protection Act 2018 and the General Data Protection Regulation (EU) 2016/679, and will be treated in the strictest confidence by the NFER, White Rose Maths and EEF. Pupil data collected from schools by NFER will not be made available to anyone outside of those parties listed. Our legal basis for gathering and using this data is legitimate interest, through our work as a research organisation.

Data protection

The legal basis for processing personal data is covered by:

GDPR Article 6 (1) (f) which states that 'processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party except where such interest

are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of the personal data’.

We have carried out a legitimate interest assessment, which demonstrates that the evaluation fulfils one of NFER’s core business purposes (undertaking research, evaluation and information activities). It has broader societal benefits and will contribute to improving the lives of learners by providing evidence for schools when making decisions about maths CPD.

In setting out the roles and responsibilities for this trial, NFER, EEF and White Rose Maths have signed a Data Sharing Agreement. This includes a description of the nature of the data being collected and how it will be shared, stored, protected and reported by each party. NFER and White Rose Maths are the joint data controllers for the trial up until the data is passed to the EEF archive.

NFER and White Rose Maths will provide a memorandum of understanding to schools, explaining the nature of the data being requested of schools and children, how it will be collected, and how it will be passed to and shared with NFER.

For the purpose of research, name, date of birth and UPN and test outcome data for all pupils in the trial will be linked with information about pupils from the National Pupil Database (held by the DfE) and other official records. Pupil and teacher data will be treated with the strictest confidence. Neither we, nor any of the named parties, will use pupil or teacher names or the name of any school in any report arising from the research.

NFER will provide the DfE Data Sharing Team at the DfE with the pupil information outlined above, allowing a match to the National Pupil database (NPD). After the matching process has taken place, NFER will then analyse this data using the Secure Research Service (SRS) based at the Office for National Statistics (ONS). NFER will access the data for analysis through the SRS secure online system. The SRS system does not allow users to remove or copy data from its servers.

At the end of EEF evaluations all data is archived to allow for further secondary analysis. At this point, EEF becomes the data controller and NFER is no longer responsible for the data and are no longer a data controller. After three months of the completion of the study, all of the matched data (i.e. to NPD) will be added to the EEF archive and ‘de-identified’ before being made available to researchers³⁴. The EEF archive is hosted by the Office for National Statistics (ONS) and managed by the EEF archive manager. Other research teams may use the de-identified data as part of subsequent research through the Approved Researcher Scheme³⁵.

We will not share personal data collected through telephone interviews or in the teacher survey with other organisations.

For further information, privacy notices for the study are available here:

<https://www.nfer.ac.uk/eej-schools-privacy-notice>

<https://www.nfer.ac.uk/eej-pupils-privacy-notice>

³⁴ De-identified means that names, dates of birth and identifiers such as unique pupil number (UPN) and teacher reference number (TRN) are removed from data before it is made available to researchers accessing the archive.

³⁵ <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme>

Personnel

Table 6: list of personnel

Name	Institute	Roles and responsibilities
Ben Styles (BS)	NFER	Project Director, responsible for leading the NFER team and project delivery.
Helen Poet (HP)	NFER	Project manager, responsible for overseeing the day to day running of the trial and the process evaluation
David Sims (DS)	NFER	Process evaluation director, responsible for overseeing the development of IPE tools
Kathryn Hurd (KH)	NFER	Test and Schools administration lead, responsible for overseeing recruitment, school contact and testing
Guido Miani (GM)	NFER	Coordinating school recruitment, school contact and testing
Joana Andrade (JA)	NFER	Statistician, responsible for statistical analysis
Eleanor Bradley (EB)	NFER	Researcher, supporting IPE
Jane Brown (JB)	White Rose Maths	Lead developer, responsible for delivery of the intervention
Tony Staneff (TF)	White Rose Maths	Lead developer, responsible for delivery of the intervention
Caroline Hamilton (CH)	White Rose Maths	Lead developer, responsible for delivery of the intervention

Risks

Table 7: List of risks and mitigations

Risk	Likelihood/ Impact	Mitigation
Insufficient schools recruited to the study	Likelihood: moderate Impact: high	Longer period for recruitment (throughout spring and summer terms, including face to face recruitment events, randomisation not until October). High level of support offered by developers and NFER throughout process. Clear and detailed recruitment materials. Clear information provided to schools explaining the principles of the trial and expectations. NFER have proven expertise in recruiting schools for randomised trials. We will provide input into the recruitment documentation and assist WRM with recruitment, if required, though a separate grant agreement.
Low school engagement/high attrition due to pressure on school from changes to statutory testing	Likelihood: moderate Impact: high	Immediate contact for data collection after MOU. Schools sign MOU with clear identification of requirements. NFER's keep in-touch activities. Use of EYFSP outcomes will result in minimal pupil attrition for the secondary outcome measures.
Intervention is not implemented well	Likelihood: low Impact: moderate	Clear information provided to schools explaining the principles of the trial and expectations. Both 'intention to treat' and CACE analysis will be used. Good communication with delivery team to aid strong implementation. Process evaluation will monitor implementation.
Researchers lost to project due to sickness, absence or staff turnover	Likelihood: moderate, especially over 3 years Impact: moderate	NFER has a large research department with numerous researchers experienced in evaluation who could be redeployed.
Further delay or disruption to delivery and/or evaluation activities due to Covid-19/school closures	Likelihood: moderate-high Impact: high	Careful monitoring and communication between all parties. We will monitor guidelines around visitors to schools at the time of the evaluation activities that would be affected (case studies, end-point test administration)

Timeline

Table 8: Timeline

Note that due to the school closures in Spring 2020, the project was paused, and subsequently it was agreed to delay it for one academic year, with the delivery and evaluation taking place from September 2021.

Dates	Activity	Staff responsible/ leading
July 2019	Project Set Up meetings	HP
November 2019	IDEA workshop	HP
October - December 2019	Protocol writing	HP/BS
November 2019 – May 2020	<i>School recruitment (original)</i>	WRM/GM/KH
March 2020	Project paused due to school closures (Covid-19)	-
November 2020 – May 2021	School recruitment – recontact schools recruited in 19/20 + top up recruitment to required numbers.	WRM/GM/KH
June/July 2021	Pupil data collection, including FSM eligibility (requested from schools after receipt of MOU)	WRM/GM/KH
Jan-June 2021	Observation of training of the trainers	WRM/HP/EB
September/October 2021	Reception Baseline Assessment Schools to complete observational checklist after their school has completed the RBA Practitioner confidence survey in schools (baseline)	GM/KH
October 2021	Randomisation	BS/JA
November 2021	Intervention commences	WRM
November 2021	First set of longitudinal case study visits (including observations)	HP/EB
January 2022	Second set of longitudinal case study interviews	HP/EB
May 2022	Third set of longitudinal case study visits Best practice case study interviews (selected by developer)	WRM/HP/EB
June/July 2022	Reception pupils sit PUMA end point test, NFER administrators in schools. Practitioner confidence survey in schools (end-point)	GM/KH/HP

Dates	Activity	Staff responsible/leading
August – October 2022	Analysis of outcomes for reception: PUMA (primary) and EYFSP (secondary) Analysis of practitioner confidence (secondary outcome)	BS/JA/HP
Oct – Dec 2022	Report writing.	HP/BS
January 2023	Peer review and provider comments.	EEF
Summer 2023	(subject to main trial security rating) NFER administrators return to schools to administer the PUMA Y1 test	GM/KH/HP
Autumn 2023	Analysis of longitudinal follow-up data (PUMA Y1) and addendum report drafted	BS/JA/HP

Appendix 1: Evidence base for Reception Jigsaw modules

Session content	Research evidence
Developing Early Number Sense <ul style="list-style-type: none"> What is number sense? The counting principles Subitising Composition of number Comparison and number relationships 	<p>*Andrews, P., Sayers, J. & Back, J. (2013) <i>The development of foundational number sense in England and Hungary: a case study comparison</i>.</p> <p>*Sayers, Andrews & Björklund Boistrup (2016) <i>The Role Of Conceptual Subitising in the Development of Foundational Number Sense</i>, Stockholm University</p> <p>EEF (2020) <i>Improving Mathematics in the Early Years and KS1</i></p> <p>Early Intervention Foundation (EIF, 2018) key competencies report: https://www.eif.org.uk/report/key-competencies-in-early-cognitive-development-things-people-numbers-and-words</p> <p>*Nicholas C. Johnson, Angela C. Turrou, Brandon G. McMillan, Mary C. Raygoza & Megan L. Franke (2019) "Can you help me count these pennies?": Surfacing preschoolers' understandings of counting, <i>Mathematical Thinking and Learning</i>, 21:4, 237-264, DOI: 10.1080/10986065.2019.1588206)</p> <p>*Sarama, J. S., & Clements, D. H. (2009). <i>Early childhood Mathematics Education Research</i>. London: Routledge.</p> <p>The Early Math Collaborative Erikson Institute (2014) <i>The big ideas of Early mathematics</i></p> <p>WW Clearing House (2013) <i>Teaching Math to Young Children</i></p>
Creating a Mathematical Classroom <ul style="list-style-type: none"> Attitudes to learning Everyday maths through classroom routines Opportunities for maths through continuous provision Role of the adult in supporting learning Planning for adult led activities Enhancing provision <p>Focus on sorting and comparison throughout</p>	<p>Bennett & Weidner (2012) <i>Everyday Maths through Everyday Provision</i></p> <p>Clements and Sarama (2009) <i>Learning and Teaching Early Math: The Learning Trajectories Approach</i>.</p> <p>Greg Bottrill (2018) <i>Can I Go and Play now?</i></p> <p>DCSF (2009) <i>Children thinking mathematically: PSRN essential knowledge for Early Years practitioners</i></p> <p>DCSF (2009) <i>Learning. Playing and Interacting</i>, National Strategies</p> <p>EEF (2020) <i>Improving Mathematics in the Early Years and KS1</i></p> <p>Early math Collaborative (2014) <i>Big Ideas of Early Maths</i></p> <p>Gifford, S (2005) <i>Teaching Mathematics 3-5</i></p> <p>Mohammed, R (2015) <i>Characteristics of Effective Learning: Play and Exploration in Action</i>.</p> <p>EY Foundation Stage Forum Article July 18 2015</p> <p>National Strategies (2009) <i>Numbers and Patterns: Laying Foundations in mathematics</i></p> <p>DCSF</p> <p>*Laevers, F (2015) <i>Making care and education more effective through wellbeing and involvement</i>. The Research Centre for Experiential Education, Belgium</p> <p>WW Clearing House (2013) <i>Teaching Math to Young Children</i></p>
Mathematical Talk and Questioning <ul style="list-style-type: none"> Why is talk important? Developing sustained shared thinking Creating opportunities for talk Open ended questions Using examples and non-examples Exploring addition and subtraction structures Developing an understanding of position 	<p>Greg Bottrill (2018) <i>Can I go and Play now?</i></p> <p>Kathy Brodie (2014) <i>Sustained Shared Thinking: Linking Theory to Practice</i></p> <p>Clements and Sarama (2009) <i>Learning and Teaching Early Math: The Learning Trajectories Approach</i>.</p> <p>Early Math Collaborative (2014) <i>Big Ideas of Early Mathematics</i></p> <p>EEF (2020) <i>Improving Mathematics in the Early Years and KS1</i></p> <p>Julie Fisher (2016) <i>Interacting or Interfering</i></p> <p>McCray et al (2019) <i>Growing Mathematical Minds</i>.</p> <p>*Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R. and Bell, D. (2002) <i>Researching Effective Pedagogy in the Early Years, (REPEY) Research Report 356</i></p> <p>*Siraj-Blatchford (2007) <i>Creativity, Communication and Collaboration: The identification of Pedagogic Progression in Sustained Shared Thinking</i></p> <p>*Sylva, Melhuish, Sammons, Siaj-Blatchford, Taggart. (2004) <i>The Effective Provision of Pre-School Education Project, funded by DFES</i></p> <p>Anne Trafton (2018) <i>Back-and-forth exchanges boost children's brain response to language MIT News Office</i></p> <p>Sir Peter Williams (2008) <i>Independent Review of Mathematics Teaching in Early Years Settings and Primary Schools Final Report</i>.</p>

Session content	Research evidence
Reasoning and Problem Solving <ul style="list-style-type: none"> • Start with provision • Start with an object • Start with a picture • Start with a game • Start with a story 	<p>Early Math Collaborative (2014) <i>Big Ideas of Early Mathematics</i></p> <p>Clements and Sarama (2009) <i>Learning and Teaching Early Math: The Learning Trajectories Approach</i>.</p> <p>DfES EYFS card (2007b) 'Learning and Development 4.1, Play and exploration'</p> <p>Woodham, L & Pennant, J (2014) <i>Mathematical Problem Solving in the Early Years</i></p>
Exploring Pattern and Shape <ul style="list-style-type: none"> • The importance of pattern spotting • Progression through pattern • Pattern through story • Numerical patterns • Odd and even • Doubling • Looking at shape 	<p>Clements and Sarama (2009) <i>Learning and Teaching Early Math: The Learning Trajectories Approach</i>.</p> <p>Erikson Early Math Collaborative (2014) <i>Big Ideas of Early Mathematics</i></p> <p>Gifford, S (2019) <i>The Case for Space in the Early Years</i> British Society for Research into Learning Mathematics</p> <p>*Gunderson, E., Ramirez, G., Beilock, S.L. & Levine, S.C. (2012). <i>The relation between spatial skill and early number knowledge: The role of the linear number line. Developmental Psychology</i>, 48(5)</p> <p>Haylock, D and Cockburn, A ((2017) <i>Understanding Mathematics for Young Children</i></p> <p>NCETM (2018) <i>Early Years Typical Progression Chart with additional guidance for practitioners – Pattern</i></p> <p>NCETM (2018) <i>Early Years Typical Progression Chart with additional guidance for practitioners – Shape and Space</i></p> <p>Montague-Smith, A, Cotton, T, Hansen, A & Price, A (2018) <i>Mathematics in early years Education</i></p> <p>Moss, J., Bruce, C.D., Caswell, B., Flynn, T. & Hawes, Z. (2016) Taking shape: activities to develop geometric and spatial thinking.</p> <p>*Rittle-Johnson B, Zippert E, Boice K (2018) <i>The Roles of Patterning and Spatial Skills in Early Mathematics Development</i></p>

* indicates that a paper tests theory of learning (information provided by WRM)