Maths Through Picture Books
Pilot Evaluation Plan

Evaluator: Oxford MeasurEd
Principal investigator: Lydia Marshall

Education Endowment Foundation

Template last updated: October 2022

PROJECT TITLE	**Pilot Evaluation of Maths Through Picture Books**
DEVELOPER (INSTITUTION)	East London Research School
EVALUATOR (INSTITUTION)	Oxford MeasurEd
PRINCIPAL INVESTIGATOR(S)	Lydia Marshall
EVALUATION PLAN AUTHOR(S)	Lydia Marshall
PUPIL AGE RANGE AND KEY STAGE	4 – 5, Reception
NUMBER OF SCHOOLS/ SETTINGS	20
NUMBER OF PUPILS	100

Evaluation plan version history

VERSION	DATE	REASON FOR REVISION
1.0	Nov 2023	
2.0	May 2024	*Appending quantitative analysis plan*

Intervention

Maths Through Picture Books is a ten-week targeted intervention that uses specific storybooks accompanied by multi-turn conversations to help children who are falling behind to develop a secure understanding of number before they move to Key Stage 1. It will be piloted with 100 children in 20 reception classes in the South East and East of England in the 2023/24 academic year.

TIDieR checklist

BRIEF NAME
Maths Through Picture Books (MTPB)

WHY
In Maths, the gap between low and high attainers in England is very wide by international standards (Ofsted, 2021). Gaps that are apparent by the end of Early Years double by the end of primary school. Most children do not catch up by the end of compulsory education and this is particularly true for disadvantaged children (EEF, 2018). There is evidence for the need to develop practitioners' understanding of how children learn mathematics, including knowledge of mathematics, children's mathematical development, and effective mathematical pedagogy (EEF, 2021a). Studies also show the significant positive impact of using teaching assistants (TAs) to provide one-to-one or small group intensive support using structured interventions (EEF, 2021b; Higgins et al., 2013; Nicoletti and Rabe, 2014).

There is a small but growing body of evidence to support the use of storybooks to teach mathematics, most of it involving the specific use of talk. The Education Endowment Foundation (EEF)'s evidence review of Early Years and Key Stage 1 Mathematics Teaching identified this as a promising approach, with six studies providing a large positive aggregate effect. All included studies conducted in early years settings involved the use of specially chosen or designed storybooks and involved guidance to support mathematical talk and educator questioning (Casey et al., 2008; Hassinger-Das et al., 2015; Purpura et al., 2017; van den Heuvel-Panhuizen et al., 2016).

The aim of the programme is to equip reception teachers and TAs with:

- improved knowledge of how children learn number and operations
- the knowledge and skills to engage a small group of targeted children in sustained multi-turn conversations about number and operations using picture books and the 'ShREC approach' (see below)

For children, the programme aims to:

- improve enjoyment of and motivation towards learning mathematics among children
- have children talk about number and operations during the intervention

WHAT
Procedures: In each school one TA and one teacher will be trained by the team from the East London Research School (ELRS). Training will involve a full-day in-person training in February 2024 and a half-day in-person training in June 2024. Training will include a focus on how children learn about number and operations, interactive reading, the EEF's Improving Mathematics in the Early Years and Key Stage 1 guidance report and embedding techniques in wider classroom practice. Additionally, during the ten-week intervention period, there will be an online session for teachers on supporting their TAs, and an online drop-in session for all teachers and TAs to support delivery. Attendance at the online session for teachers will be

recorded by ELRS and teachers directed to watch the recording of the session if they are unable to attend.

TAs run 20-minute sessions doing dialogic book reading with a set of nine specially chosen maths-focused picture books, supported by semi- structured prompts for the TA. Conversations are structured using the ShREC framework. The aim of the ShREC framework is to provide early years practitioners with a simple and memorable set of evidence-informed strategies that can be embedded into everyday practice (EEF 2022b). This includes:

- **Sh**aring attention – being at the child's level and paying attention to what they are focused on
- **R**espond – following the child's lead and responding to their non-verbal and verbal communications
- **E**xpand – repeating what the child says and building on it by adding more words to turn it into sentences
- **C**onversation – having extended back and forth interactions and giving children time to listen, process and reply

The TA and teacher will have weekly planning meetings to ensure fidelity to the programme and consistency between the targeted sessions and wider classroom practice. At each planning meeting, the TA and teacher will reflect on the previous week's sessions, recording children's attendance, any challenges, and child-specific reflections. They will then prepare for the following week's sessions by looking through the new book, reading the scaffold and discussing the sessions in advance. This will include identifying the mathematical focus of the sessions and using the ShREC framework to structure potential conversations, as well as discussing what the TA will be intentionally focusing on for each child. Finally, they will think through links to the classroom and opportunities to revisit, reinforce and rehearse the mathematical focus in wider practice. The teacher will model the sessions if required. This will all be recorded on a planning template that will be uploaded to the school's secure Google Drive. The TA and teacher have the option of a) typing the information directly into the planning template in the secure Google Drive or b) printing out a planning template, filling it in by hand, then scanning and uploading it to the Google Drive.

During delivery, teachers will observe the TA delivering the programme three times (in Weeks 1, 4 and 10). Teachers will use the ShREC framework while they are observing the session and make notes in an observation template. The teacher will have a professional conversation with the TA on the same day, providing feedback and supporting the TA to reflect on the session using the ShREC framework. Teachers will use techniques that they have been taught in the module on implementation and support for colleagues. Their joint reflections will be recorded on the observation template and uploaded into the school's secure Google Drive.

Materials: Schools are provided with resources to support delivery and two sets of nine books for classroom use (one for the intervention and one for use in the classroom outside of the intervention). Resources include a programme handbook, planning template, observation template, ShREC framework, video exemplification and the picture book scaffolds. All resources, templates and frameworks will be found in the handbook and on Padlet.

WHO PROVIDED

ELRS will provide the training and support for schools. Fliss James and Melissa Prendergast are highly experienced early years teachers and leaders. They have significant expertise in designing and delivering high quality evidence-informed professional development. TAs will deliver the in-school intervention sessions, supported by teachers.

HOW

The intervention is delivered face-to-face to groups of five children who teachers have identified as requiring extra support. To select children for the pilot, the participating teacher will carry out a short activity based on the Give-N task (see Appendix 3),[1] with all children in the reception class. This activity will enable the teachers to identify children who have not yet developed a secure understanding of cardinality, i.e., being able to count numbers in a set. Teachers are provided with a Padlet which contains guidance on how to carry out the activity and an observation grid for recording the key information (see Appendix 3). Where teachers identify more than five children who find the task challenging, they are encouraged to use their professional judgement and knowledge of each individual child to select those children they believe would benefit most from the programme.

WHERE

MTPB sessions run in a dedicated, quiet space away from the classroom. The participating schools need to be big enough to support this requirement.

To enable the in-person training and to cover a diverse sample, the delivery will be in primary schools in London Boroughs of Redbridge, Havering, and Newham and in Essex.

WHEN and HOW MUCH

MTPB sessions are delivered twice per week for ten weeks. The sessions will last for 20 minutes and will happen at the same time every week.

TAILORING

The intervention should be delivered as intended without any significant adaptations being made.

MODIFICATIONS

During a formative evaluation of the programme's early-stage development work, some alterations were found to be made in classrooms, such as including manipulatives[2] in the small group reading sessions, that have highlighted the need to be more explicit in the communications around acceptable adaptation by the delivery team. Some elements of the programme are flexible, such as the day the sessions are held, the addition of extra sessions to share the same book and sending the book home or adding it to the class library. The evaluation will monitor any adaptations being made.

HOW WELL

In the first training session, ELRS will systematically talk though the programme using a simple logic model, ensuring that participants have a clear understanding. Participants will understand where there is room for them to make adaptations and the aspects of the programme that must always be present. The handbook and Padlet will provide step by step instructions and support so that schools can run the programme with fidelity. ELRS will monitor the fidelity with which the programme is delivered each week using the planning templates that TAs and teachers will be submitting. This will enable ELRS to identify practice that is in line with expectations, identify barriers, and provide support as appropriate.

[1] The Give-N task is widely used in the field of Developmental Psychology to indicate young children's knowledge or use of the cardinality principle, wherein the last number word used in the counting process indicates the total number of items in a collection (see Baroody and Lay 2022).

[2] Mathematical manipulatives are artifacts used in mathematics education, enabling children to explore or investigate mathematical concepts or processes and to perform problem-solving activities. An example is plastic blocks that click together, often used to learn addition and subtraction. Images of these are included in one of the storybooks and practitioners have improvised to include real-life blocks while conversing with children.

Existing evidence

A formative evaluation of the programme's early-stage development work suggested that the programme is practical to implement and has high acceptability. Teachers and TAs enjoyed delivering the programme and were highly appreciative of the professional development element of the programme. The programme was a positive experience for children who were reported to have built confidence and enjoyment in maths, with data suggesting that most of the children made considerable progress in a short amount of time. The evaluation concluded that it was implemented with a high degree of fidelity, however there were some alterations made in the classroom (as stated above). Delivery in some schools was impacted by staff and student absence, with the timetabling of sessions into the busy weeks of the summer term proving a challenge. It was shown to be easier to implement in larger schools due to staff availability.

Theory of Change

The programme's Theory of Change (ToC) and log of associated causal and contextual assumptions are included in Appendix 1 of this study plan.

Research questions

The evaluation will address 15 research questions, mapped in Table 1 below against the EEF's three pilot criteria – evidence of promise, feasibility of implementation and readiness for trial (including scalability).

The research questions have been developed in consultation with the EEF and the delivery team during the project inception phase. The research questions about evidence of promise address the intended and unintended consequences stated in the ToC. The research questions about feasibility reflect the causal mechanisms and contextual assumptions set out in the ToC and associated logs. The research questions about readiness for trial consider scalability and the ability of an evaluation to test the programme's ToC, including intended impacts.

Table 1 Research questions

Pilot criteria	Research questions
Evidence of promise	1. Is there an improved understanding among TAs and Teachers, of how children learn about numbers and operations?
	2. Is there an increased confidence in teaching maths among TAs and Teachers?
	3. Do children talk about number and demonstrate increased enjoyment of and motivation towards learning mathematics? • Are disadvantaged children perceived to benefit more or less than their peers?
	4. Is teaching practice embedded and sustained after 10 weeks? • Is the use of picture books to teach maths sustained? • Is the use of principles around selecting picture books to teach maths sustained? • Is the use of ShREC strategies sustained?
	5. Are there any unintended consequences of participation in this programme? • Are ShREC strategies used outside of the intervention? • Is the intervention perceived to have benefits for children's socioemotional learning and/or language and communication? • Are there benefits for child-adult relationships? • Does the intervention take practitioners away from other activities and is this trade-off perceived to be worth it? • Does the intervention take children away from other activities and is this trade-off perceived to be worth it? • Do children selected for the intervention feel stigmatised? • Do children not selected feel they are missing out? • Do any unintended consequences affect disadvantaged pupils more or less than their peers?
	6. How different is the programme to business as usual?
Feasibility of implementation	7. Is the programme delivered as intended? • Is training and support for practitioners delivered as intended? • Is the programme delivered as intended in school?
	8. What contextual factors affect implementation? • Are there any particular challenges around including disadvantaged children?
	9. What adaptations are participants making to the programme and why?
	10. How appropriate are the pupil selection/eligibility criteria? • How well do the selection/eligibility criteria enable schools to include disadvantaged children? • How might the timing of selecting pupils pre-testing and pre-training influence intervention outcomes?
	11. Are programme inputs (training sessions, physical resources, additional support) accessible and useful to practitioners? • How well do the inputs support practitioners to include and support disadvantaged children?
	12. Is the programme affordable for schools?[3]
Readiness for trial	13. Can the programme feasibly be delivered at a larger scale? • Is there a clear and realistic plan for what scaling to deliver in more schools would look like and require? • What modifications would be required to scale delivery?

[3] Please see Table 5 for more details about questions related to programme costs.

	• How might the timing of selecting pupils pre-testing and pre-training influence intervention outcomes?
	14. How well do we understand the causal chain and mechanisms which underpin the intended effects of the programme?
	15. What potential measures could be used in a subsequent efficacy trial?

We will report on 15 key success indicators to inform the decision about whether MTPB should proceed to trial. Table 2 maps these success indicators against the three pilot criteria and the research questions listed above and provides the sources of evidence for each indicator and how they will be assessed. The sources of evidence are all explained under Methods below.

Table 2 Success indicators

Pilot criteria	Success indicators	Sources of evidence
Evidence of promise	**Improved understanding among TAs and teachers of how children learn about numbers and operations (RQ1)**	• Summer (post-delivery and Autumn term (longitudinal follow-up) surveys (quantitative teacher/TA reported improvements) • Qualitative interviews with teachers/TAs (qualitative self-reported improvements – the how and why)
	Increased confidence in teaching maths among TAs and teachers (RQ2)	• Summer and Autumn term surveys (quantitative teacher/TA-reported increases) • Qualitative interviews with teachers/TAs (qualitative self-reported increases – the how and why)
	Children talk about number and operations during the intervention (RQ3)	• Summer and Autumn term surveys (quantitative TA reports on proportion/regularity of children talking about number/operations) • Qualitative interviews (qualitative TA reports on the how and why) • Observations (qualitative and quantitative observations)
	Improved enjoyment of and motivation towards learning mathematics among children (RQ3)	• Summer and Autumn term surveys (quantitative teacher/TA reports on improvements) • Qualitative interviews (qualitative teacher/TA reports on improvements – the how and why) • Child discussion groups (qualitative self-reported improvements – the how and why)
	Teaching practice embedded and sustained after 10 weeks (RQ4)	• Autumn term survey (quantitative teacher/TA-reported practices)

Pilot criteria	Success indicators	Sources of evidence
	No indication that the programme likely widens the gap between disadvantaged pupils and their peers (RQ5)	• Qualitative interviews with a member of the Senior Leadership Team (qualitative perceptions of promise) • Qualitative interviews with teachers/TAs (qualitative perceptions of promise)
Feasibility of implementation	**Programme is delivered with fidelity (RQ7)**	• Training attendance data • Other monitoring information from Padlet • Summer term survey (quantitative teacher/TA-reported fidelity) • Qualitative interviews with teachers/TAs (qualitative self-reported fidelity and adaptation – the how and why) • Observations (qualitative and quantitative observations of agreed core components)
	Pupil selection/eligibility criteria are appropriate and feasible to implement (RQ10)	• Qualitative interviews with teachers/TAs (qualitative perceptions of accessibility and usefulness) • Summer term survey (quantitative teacher/TA perceptions)
	Programme inputs (training, physical resources, additional support) are accessible and useful (RQ11)	• Qualitative interviews with teachers/TAs (qualitative perceptions of appropriateness and feasibility) • Summer and Autumn term surveys (quantitative teacher/TA perceptions)
	Programme is affordable to schools (RQ12)	• Qualitative interviews with a member of the Senior Leadership Team (qualitative perceptions of affordability) • Qualitative interviews with teachers/TAs (qualitative perceptions of affordability of demands on time) • Summer term survey (quantitative teacher/TA-reported costs)
Readiness for trial	**Core programme elements identified (RQ13)**	• IDEA workshop • Desk-based review • Learning workshop

Pilot criteria	Success indicators	Sources of evidence
	Clear and realistic plan for what scaling to deliver in more schools would look like and require (RQ13)	• Interviews with delivery team • Learning workshop
	Plan for scaling would entail little or no modification to in-school delivery (RQ13)	• Interviews with delivery team • Learning workshop
	Causal chain and underpinning mechanisms articulated (RQ14)	• IDEA workshop • Desk-based review • Learning workshop
	Outcome measure(s) identified that is/are **a) aligned to intended outcomes,** **b) able to measure changes that we expect to see and** **c) acceptable to schools (RQ15)**	• Desk-based review • Piloting assessments [primary measure for a trial] • Child discussion groups [secondary measure for a trial]

Methods

Our design includes desk-based research, primary research with the delivery team (DT), teachers, TAs and school Senior Leadership Team members (SLT) and children, and observations of MTPB delivery.

Figure 1 sets out how our design will address the three pilot criteria and 15 research questions.

Figure 1 Evaluation matrix

Evaluation domain	Research question	Desk-based research			Research with DT			Research with school staff					Observations		Research with pupils	
								Interviews			Surveys					
		Programme documentation	Programmatic data	Outcome measure review	IDEA workshop	Interviews with delivery staff	Learning workshop	TAs	Teachers	SLT	Teachers	TAs	Training	Classroom	Outcome measure piloting	Pupil focus groups
Evidence of promise	1. Is there an improved understanding, among TAs and Teachers, of how children learn about numbers and operations?							■	■		■	■				
	2. Is there an increased confidence in teaching maths among TAs and Teachers?							■	■	■	■	■				
	3. Do pupils talk about number and demonstrate increased enjoyment of and motivation towards learning mathematics?							■	■							■
	4. Is teaching practice embedded and sustained after 10 weeks?										■	■				
	5. Are there any unintended consequences of participation in this programme?					■	■	■	■	■	■	■		■		■
	6. How different is the programme to business as usual?						■	■	■	■	■	■				
Feasibility	7. Is the programme delivered as intended?	■				■	■	■	■	■	■	■	■	■		
	8. What contextual factors affect implementation?	■	■			■	■	■	■	■	■	■		■		
	9. What adaptations are participants making to the programme and why?		■			■		■	■	■	■	■		■		
	10. How appropriate are the pupil selection/eligibility criteria?							■	■	■	■	■				
	11. Are programme inputs (training sessions, physical resources, additional support) acceptable and appropriate?					■		■	■	■	■	■		■		■
	12. Is the programme affordable for schools?							■	■	■	■	■				
Readiness for trial	13. Can the programme feasibly be delivered at a larger scale?					■	■									
	14. How well do we understand the causal chain and mechanisms which underpin the intended effects of the programme?				■		■									
	15. What potential measures could be used in a subsequent efficacy trial?			■											■	■

Recruitment

The delivery team will recruit 20 schools for the pilot evaluation. The chosen sample size considered both ELRS's capacity constraints, that allow them to deliver training and support to no more than 25 schools, and evaluation requirements, which requires a minimum of 16 schools to enable 8 – 10 evaluation interviews to be conducted without both teacher and TA at any one school being interviewed. Oxford MeasurEd, ELRS and the EEF decided on 20 as the sample size to deliver best value for money.

The delivery team and the evaluation team will monitor school recruitment to ensure that:

- participating schools are not participating in other EEF-funded maths/numeracy focused interventions in Reception, to avoid contamination
- we achieve a diverse sample, including a mix of rural and urban schools

Participating schools will nominate a TA and teacher from the same reception class to take part in the intervention when they sign up to the pilot. There will be no eligibility criteria for these practitioners, other than the school authorising their time to be spent attending training and delivering the intervention. The teacher, TA and head teacher will sign a Memorandum of Understanding (MOU) agreeing to the requirements of the programme and the evaluation.

To select five children for the pilot, the participating teacher at each school will carry out a short activity based on the Give-N task, with all children in the reception class. Where schools have multiple reception classes, the activity will only take place with children in one reception class. The activity will be done in the second half of the Autumn term, as children will have settled into reception by then. This activity will enable the teachers to identify children who have not yet developed a secure understanding of cardinality, i.e., being able to count numbers in a set. Teachers are provided with a Padlet which contains guidance on how to carry out the activity and an observation grid for recording the key information (see Appendix 3). Where teachers identify more than five children who find the task challenging, they are encouraged to use their professional judgement and knowledge of each individual child to select those children they believe would benefit most from the programme.

To prevent bias where the project delivery team will be involved, we (the evaluation team at Oxford MeasurEd) will prepare all recruitment materials for specific evaluation activities. These recruitment materials will outline the purpose of the evaluation, what will be involved in taking part and how data will be used. We will also explain this verbally before all workshops, interviews and observations, and participants will be able to ask questions before agreeing to take part. Recruitment documents and verbal briefings will also make clear the steps we will take to minimise the chance of participants being identifiable in reports.

Measures

Short-term outcomes

The short-term outcomes stated in the programme's ToC (included in full in Appendix 1) include:

For practitioners:

- improved understanding of how children learn about numbers and operations
- improved understanding of how to use picture books interactively (using the ShREC approach) to encourage mathematical conversations
- teachers supporting TAs to develop their professional expertise

For children:

- talking about number and operations during the intervention

- improved enjoyment of and motivation towards learning mathematics

As set out in Table 2 above, these outcomes will be measured as follows:

Table 3 Evidencing short-term outcomes

Success indicators	Sources of evidence
Improved understanding among TAs and teachers of how children learn about numbers and operations (RQ1)	• Summer and Autumn term surveys (quantitative teacher/TA-reported improvements) • Qualitative interviews with teachers/TAs (qualitative self-reported improvements – the how and why)
Children talk about number and operations during the intervention (RQ3)	• Summer and Autumn term surveys (quantitative TA reports on proportion/regularity of children talking about number/operations) • Qualitative interviews (qualitative TA reports on the how and why) • Observations (qualitative and quantitative observations)
Improved enjoyment of and motivation towards learning mathematics among children (RQ3)	• Summer and Autumn term surveys (quantitative teacher/TA reports on improvements) • Qualitative interviews (qualitative teacher/TA reports on improvements – the how and why) • Child discussion groups (qualitative self-reported improvements – the how and why)

Longer-term outcomes

The longer-term outcomes articulated in the programme's ToC include:

For practitioners:

- improved confidence in teaching number and operations
- embedding the ShREC approach leading to high quality mathematical talk in everyday practice
- improved ability to make contingent responses when using mathematical picture books interactively
- teachers having improved skills in supporting and developing TAs to deliver evidence-informed interventions

For children:

- improved understanding of number and operations
- improved Key Stage 1 maths attainment

This pilot is not designed to measure long-term impact. However, as set out in Table 2, it will collect some data on longer-term practitioner outcomes, as follows:

Table 4 Evidencing longer-term outcomes

Success indicators	Sources of evidence
Increased confidence in teaching maths among TAs and teachers (RQ2)	• Summer and Autumn term surveys (quantitative teacher/TA-reported increases) • Qualitative interviews with teachers/TAs (qualitative self-reported increases – the how and why)

Success indicators	Sources of evidence
Teaching practice embedded and sustained after 10 weeks (RQ4)	• Autumn term survey (quantitative teacher/TA-reported practices)

Attainment measures

Data from all children participating in the pilot will be collected using the Pearson Wechsler Individual Achievement Test - Third UK Edition (WIAT-III UK). The WIAT-III UK comprises several subtests. For the purpose of this evaluation, we will administer the Math Problem Solving subtest at baseline, which is suitable for children aged 4 and above. We are not administering the Numeracy subset at baseline as it is only suitable for children aged 5 and above. At endline, however, we will administer the Numeracy subtest alongside the Math Problem Solving subtest.

All participating pupils will take part in order to test the procedures that would be used in a trial and to maximise the sample available for generating statistical information to inform future trial designs. These measures will **NOT** be used to assess impact.

Data collection

Desk-based research

Before finalising the evaluation design, we ***reviewed programme documentation*** to:

- understand the ToC, including causal mechanisms and contextual assumptions
- map research evidence collected against the causal mechanisms and contextual assumptions so far
- identify any programmatic data that we might use for the evaluation

We also carried out a desk-based ***review of potential outcome measures*** for a future trial. We agreed the methodology for this review with the EEF in April 2023. Our review included considering, for each potential outcome measure:

- the alignment between what they measure and the intended outcomes of MTPB
- their statistical properties
- logistical factors such as mode of data collection, length of assessment and any training/briefing required to administer the assessment
- their suitability for use with reception children

We analysed the pros and cons of each identified measure and outlined these for discussion with the EEF and delivery team. Following the review, we identified the WIAT-III UK as an appropriate measure for testing children participating in the MTPB programme. The WIAT-III UK is a comprehensive assessment of attainment administered orally to each child individually. This measure was chosen as it is suitable for use with reception-aged children, has been designed for use in the UK and has been UK norms tested using a sample size of 3,000, has good construct validity and internal consistency, and requires only 10 minutes to administer each subtest. This will be piloted with children in this evaluation (see below). A note outlining the methodology and findings of the review have been included in Appendix 2 of this study plan.

We will also complete a light touch review of ***other early years maths provision in the pilot areas*** in order to understand business as usual (BAU) and the context in which pilot schools are implementing the intervention. As well as desk-based research this may include conversations with key informants (e.g., Maths Hubs or local authority representatives) in each of the areas.

Finally, we will ***analyse programmatic data*** to understand how MTPB was delivered, assess fidelity, and explore variation/adaptation. This analysis will include quantitative measures of dosage (attendance at training, completion of logs in Padlet) and light touch analysis of *how* TAs and teachers use the online platform. It will not include analysis of the content of logs in order to assess evidence of promise in terms of change in practitioner or child outcomes.

Research with delivery team

The delivery team will be the first participant group for our primary research, as well as a key audience for our findings.

We facilitated an in-person ***IDEA workshop*** with the delivery team in February 2023, following EEF guidance. In this workshop we discussed the ToC, assumed causal mechanisms and contextual assumptions, defined compliance and fidelity and identified priorities for the pilot evaluation to explore in terms of fidelity and adaptation.

We will carry out ***qualitative interviews*** with members of the team who deliver the MTPB training. These interviews will be conducted to gather their views on the feasibility of the training, the differing training needs of teachers versus TAs, evidence of promise in terms of TA and teacher understanding of the MTPB methods, and reflections on scalability.

Finally, we will facilitate a ***post-delivery learning workshop*** with the delivery team.

In this in-person workshop we will:

- explore lessons learned by the team: understanding any challenges encountered and how they were overcome and identifying any adaptations that they would intend to make in the future
- discuss the delivery team's plans for scaling delivery: exploring what scaling delivery to the volume required for an efficacy trial would entail, any modifications that would be required and any anticipated challenges
- seek participant validation of our analysis: presenting emerging findings from our evaluation for the delivery team's input to inform our final analysis and reporting

Research with school staff

We will conduct quantitative and qualitative research with school staff to achieve depth and breadth in our understanding of their experiences and views. This will include:

- **TAs:** responsible for delivering MTPB in classrooms; take part in MTPB training, of which they are the key beneficiaries
- **Teachers:** responsible for supporting quality and fidelity of the TA's delivery of MTPB and ensuring MTPB delivery is consistent with wider classroom practice; take part in MTPB training
- **SLT members**: provide a school-level perspective on MTPB

Qualitative interviews with school staff

We will conduct ***qualitative interviews*** during the 10-week intervention period with TAs, teachers and SLT members between March and May 2024. Table 5 sets out indicative topics for each participant group.

Table 5 Topics for interviews with school staff

Interviews with TAs	Interviews with teachers	Interviews with SLT
Evidence of promise: • Self-reported practitioner-level outcomes • Perceived pupil-level outcomes, including for different groups • Perceived mechanisms of change • Intentions to sustain practice • Unintended consequences • Views on programme's difference to BAU **Feasibility:** • Accessibility and usefulness of training and support from ELRS • Accessibility and usefulness of training and support from teacher • Appropriateness and feasibility of pupil selection criteria • Fidelity of delivery in school • Adaptations and reasons for these • Barriers and facilitators to successful delivery, including ShREC strategies and including for different groups of children • Affordability of demands on TA time **Readiness for trial:** • Recommendations for future delivery	**Evidence of promise:** • Self-reported practitioner-level outcomes • Perceived practitioner-level outcomes for TA • Perceived pupil-level outcomes, including for different groups • Perceived mechanisms of change • Intentions to sustain practice • Unintended consequences • Views on programme's difference to BAU **Feasibility:** • Accessibility and usefulness of training and support from ELRS • Appropriateness and feasibility of pupil selection criteria • Fidelity of delivery in school, including support to TA • Adaptations and reasons for these • Barriers and facilitators to successful delivery, including ShREC strategies and including for different groups of children • Affordability of demands on teacher and TA time **Readiness for trial** • Recommendations for future delivery	**Evidence of promise:** • Perceived practitioner-level outcomes for TA and teacher • Intentions to sustain practice • Unintended consequences • Views on programme's difference to BAU **Feasibility:** • Feasibility of incorporating MTPB within the school's existing activities • Barriers and facilitators to participation and successful delivery • Affordability of demands on teacher and TA time • Affordability of potential programme costs **Readiness for trial** • Recommendations for future delivery

Quantitative surveys of school staff

We will invite all TAs and teachers to complete a *short,* ***online post-delivery survey*** shortly after the end of delivery to quantify their experiences and views on the programme's evidence of promise and feasibility, and perceptions of benefits and unintended consequences. The survey will also ask about intentions to continue with the MTPB intervention and/or with using the ShREC strategies or picture books to teach children about number, and about other early years maths interventions that the school engaged with that year.

We will also carry out a second, **online *longitudinal follow-up survey*** of TAs and teachers in the first term of the academic year following the pilot to capture self-reported sustained practice.

The questions in each survey will be tailored to the roles of implementer (TAs) and supervisor (teachers), as set out in Table 6.

Table 6 Topics for surveys of school staff

Timing	TAs	Teachers
Post-intervention (Summer term)	**Evidence of promise:** • Self-reported practitioner-level outcomes • Perceived pupil-level outcomes • Intentions to sustain practice • Unintended consequences • Other early years maths interventions • Views on programme's difference to BAU **Feasibility:** • Usefulness of training and support from delivery team • Usefulness of support from teacher • Fidelity of delivery in school • Barriers and facilitators to successful delivery • Time spent on the intervention and any financial costs	**Evidence of promise:** • Self-reported practitioner-level outcomes • Perceived practitioner-level outcomes for TA • Perceived pupil-level outcomes • Intentions to sustain practice • Unintended consequences • Other early years maths interventions • Views on programme's difference to BAU **Feasibility:** • Usefulness of training and support from delivery team • Fidelity of delivery in school, including support to TA • Barriers and facilitators to successful delivery • Time spent on the intervention and any financial costs
Longitudinal follow-up (Autumn term)	**Evidence of promise:** • Sustained practice • Self-reported sustained practitioner-level outcomes • Perceived sustained pupil-level outcomes	**Evidence of promise:** • Sustained practice • Self-reported sustained practitioner-level outcomes • Perceived sustained practitioner-level outcomes for TA • Perceived sustained pupil-level outcomes

Training and targeted session observations

We will observe the in-person and online training sessions for TAs and Teachers in February 2024. We will also observe a sample of TA-led sessions with children and – where possible – the accompanying Teacher-TA reflection session (taking place in Weeks 1, 4 and 10) or planning session (taking place on a weekly basis). All in-school observations will take place between March and May 2024.

The primary aim of the observations will be to provide insights to tailor the prompts and probes used in interviews with the delivery team, teachers and TAs. We will also gather observational data on:

- teachers and TAs' responsiveness to training
- children's responsiveness to MTPB sessions
- any challenges to successful delivery and strategies to overcome these
- any adaptations to delivery

For the first two observations in Week 1, two researchers will conduct the observation together to ensure consistent usage of the observation pro-forma for the remaining observations.

Research with children

Our research with children will include qualitative research to understand their perspectives on evidence of promise and feasibility, and piloting of outcome measure data collection to inform our understanding of readiness for trial.

Qualitative research with children

MTPB aims to improve children's enjoyment and motivation towards learning mathematics. To include children's perspectives, we will carry out child discussion groups using age-appropriate facilitation techniques to explore children's:

- feelings about being selected for the intervention
- views on the session content
- enjoyment of/motivation towards learning mathematics

Two researchers, including a senior team member, will attend the first two group discussions with children to ensure quality assurance and ensure consistency in data collection.

Our approach draws on good practice guidance (see for example Brady and Graham 2019) and includes key considerations such as:

- keeping activities as brief as possible to maintain focus
- being flexible and planning a range of alternative activities in advance
- using projective techniques such drawing and using pictures to elicit children's feelings about taking part in MTPB and about maths in general
- engaging small groups in activities together as well as circulating to speak quietly to children one-on-one about their thoughts and feelings
- using age-appropriate language, e.g., talking about numbers and about the books they look at with [named TA], rather than asking for example how much they enjoy maths

During the process of drafting the data collection tools, we will consult closely with ELRS as experts in both the intervention and in working directly with this age group.

Carrying out meaningful research with children of this age is a challenging endeavour. These discussion groups will also serve as a pilot for how pupil voices might be included in a future trial.

Piloting of outcome measure data collection

We will collect baseline and endline outcome data for all children participating in the pilot using the WIAT-III UK, the assessment we identified in our desk-based review of outcome measures. We will collect outcome data using the WIAT-III UK Math Problem Solving subtest at baseline and outcome data from the WIAT-III UK Math Problem Solving and Numeracy subtests at endline.

The assessment will be administered digitally using tablets by a team of externally sourced teachers managed by Oxford MeasurEd, to reduce the burden of the evaluation on the schools. The assessment is administered one-to-one with the child and each subtest is estimated to take around 10 minutes to complete. We therefore estimate around 20 minutes total at baseline and 30 minutes at endline per child to include the time required to administer the test and for the child to get settled before the assessment.

A member of the Oxford MeasurEd evaluation will carry out "spot checks" at approximately 20% of schools at baseline and endline to ensure that the assessments are carried out as intended and to a high standard. We will also carry our remote quality assurance: reviewing

assessment data on a daily basis and following up with the teacher-assessors about any missing data or inconsistencies in the data.

The purpose of this piloting will be twofold:

- to assess the feasibility and acceptability of processes that would be used to collect outcome data in a trial (i.e., pupil enumeration, baseline and endline testing, allowing an external adult into school to carry out assessments)
- to gather key statistical information that can inform the design of a future trial (e.g., pre-post correlation, ICCs, data missingness)

As well as collecting the outcome data, we will carry out a short interview with responsible school staff and test administrators in a small number of schools to understand the acceptability of the process and how it might be improved, identify any key risks and challenges, and explore how schools might best be incentivised to use this measure.

Sampling approach

Table 7 sets out sampling strategies and proposed sample sizes.

For the qualitative research with schools, we would aim to interview *either* the TA *or* the teacher in each pilot school. This approach has the advantage of:

- including the experiences of all schools involved in the pilot
- reducing burden on schools
- enabling teachers and especially TAs to speak more freely about their experiences

We will carry out more in-depth research with 8-10 schools, where we will also observe delivery and interview a relevant SLT member (for example, Head of Reception/Early Years Foundation Stage or Maths Lead). In five we will also carry out a child discussion group, to which we will invite all children taking part in Maths Through Picture Books. For these schools, we will be able to offer the option of in-person as well as remote online interviews with staff.

We will collect background data (demographics and qualifications) about teachers and TAs who take part in the pilot when they join the evaluation. We will use this information to consider whether it is desirable and feasible to sample teachers and TAs purposively for interviews and schools purposively for the more in-depth research. For example, if there is a diversity of gender, ethnicity, qualification levels or years of experience in the participating population of teachers and/or TAs, we may be able to set out to include teachers and TAs with a range of these characteristics in our sample. However, we will be limited in the extent to which we can select participants due to a) our small sample and b) our decision to include all participating schools in the qualitative research. If we can sample purposively, we will agree this strategy in advance with the EEF. Either way, we will describe in the final evaluation report the characteristics of the population participating in the pilot and our final achieved samples for the interviews and surveys and highlight any possible issues with practitioners with certain characteristics being under-represented.

Table 7 Sampling

Participant group	Activity	Sampling strategy	Timing	Mode and duration	n.
Delivery team	**IDEA workshop**	Strategic stakeholders	Inception	In-person, 2 hours	1 workshop
	Qualitative interviews	Delivery team members	During 10-week intervention	Online, c. 45 mins	c. 2 (tbc)
	Learning workshop	Strategic stakeholders	Post-delivery	In-person, 2 hours	1 workshop
TAs	**Qualitative interviews**	Purposive sample	During 10-week intervention	In-person or online, c. 45 mins	10
	Post-intervention survey	Census of all TAs	Post-delivery (Summer term)	Online, c 15 mins	c. 20
	Longitudinal follow-up survey	Census of all TAs	Post-delivery (Autumn term)	Online, c. 5 mins	c. 20
Teachers	**Qualitative interviews**	Purposive sample	During 10-week intervention	In-person or online, c. 45 mins	10
	Post-intervention survey	Census of all teachers	Post-delivery (Summer term)	Online, max 20 mins	c. 20
	Longitudinal follow-up survey	Census of all teachers	Post-delivery (Autumn term)	Online, c. 5 mins	c. 20
SLT	**Qualitative interviews**	Purposive sample	During 10-week intervention	In-person or online, c. 45 mins	8-10
Observations	**Training**	Census of all training and webinars	During 10-week intervention	In-person and online as per training mode	2 in-person days 3 webinars
	In-school delivery	Purposive sample	During 10-week Intervention	In-person	8-10
Children	**Outcome measure piloting**	Census of all children	January 2023 (baseline) June-July 2023 (endline)	In-person, c. 10 (baseline) to 20 (endline) minutes	c. 100 children in 20 schools
	Child discussion groups	Purposive sample	Towards end of delivery	In-person, c. 25 mins	5 groups of 5 children

Data analysis

We will record workshops and interviews with participants' permission and use the 'Framework' approach to manage qualitative data and carry out within and cross-case analysis (Ritchie et al., 2013). Using the themes covered in topic guides, we will assemble a matrix in which each row represents an individual interview and each column a theme, and summarise data in the matrix, including illustrative quotes where appropriate. Once we have coded all

data in this matrix, we will analyse it thematically. This will include applying the themes from topic guides that the matrix is structured by and identifying new themes emerging from the data, meaning that the analysis is in part deductive, and in part inductive. We will be able to carry out both descriptive and explanatory analysis, identifying convergence and dissonance between individuals and between participant groups and looking for explanations for differing views.

We will analyse quantitative survey and programmatic data descriptively using unweighted frequencies and correlation or 'cross-tab' analysis in Stata. We have agreed a detailed analysis plan with the EEF, which can be found in Appendix 5: Survey analysis plan. The Evaluation Director will quality assure the analysis syntax and output, including confirming consistency with the analysis plan and carrying out spot checks to check for any errors or omissions in the analysis.

We will analyse quantitative data from the outcome measure piloting to provide key statistical information that can inform the design of a future trial (e.g., pre-post correlation, ICCs, data missingness). The statistical information will be analysed in Stata 18 and will include the following:

- Assessment of the pre- and post- outcome distribution to assess for floor and ceiling effects through histograms.
- Reporting the mean, standard deviation and median of the measure at baseline and endline.
- Estimation of the Pearson correlation between pre- and post- outcomes.
- An estimation of how much variance in the post- outcome is explained (R^2) by the pre-test using multi-level linear regression model. The dependent variable will be the post-outcome, regressed against the pre- outcome, with a random intercept for providers. A full regression formula is provided in Appendix 4: Outcome analysis plan.
- The ICC will be estimated from the multi-level regression model with the postestimation command **estat icc.** The formula for the ICC estimation is provided in Appendix 4: Outcome analysis plan.
- Attrition will be assessed by the proportion of pupils who completed a pre-test who did not take part in post- outcome testing.
- Item missingness will be assessed through tabulations.

We will not analyse the data to measure changes in pupil outcomes over the course of the pilot since a) we have a small sample size (max n. = 100) and b) we do not have a counterfactual estimate of the progress children would make in the absence of the intervention. Our Impact Evaluation Expert will quality assure the outcome measure analysis syntax and output.

Our analysis will be ongoing throughout the pilot. This ensures that we can provide formative outputs and follow up on emerging findings in later research activities wherever possible. Throughout, we will triangulate quantitative and qualitative findings to address the three evaluation domains and final evaluation questions. This triangulation will culminate in an internal analysis workshop, led by our Evaluation Director. The evaluation team will come together to discuss qualitative and quantitative findings from different sources under each of the research questions and identify convergent and divergent findings, which we will explain in our reporting. This form of peer discussion and review for mixed method research also serves the purpose of quality assurance, as it ensures legitimacy of findings (EEF, 2022a).

Where findings diverge between sources, we will revisit the raw data and consider potential theoretical and methodological reasons for these differences.

Our reporting process will also include participant validation at a learning workshop with the delivery team. At the learning workshop, we will present the emerging findings and seek reflections from the delivery team to contextualise and enrich our final conclusions and recommendations.

Ethics and registration

The evaluation was reviewed and approved by the Oxford MeasurEd ethics board in August 2023.

We are committed to the highest ethical standards. The project team will be guided by UK Evaluation Society Guidelines for Good Practice and comply with our Code of Conduct, Safeguarding, and Anti-Bribery and Corruption policies. Some key ethical considerations for this project will include:

- Obtaining **informed consent**: We will provide recruitment materials outlining the purpose of the evaluation, what will be involved in taking part and how data will be used. We will also explain this verbally before all workshops, interviews and observations and participants will be able to ask questions before agreeing to take part.

- Ensuring participation is **voluntary**: Schools will sign an MOU agreeing to take part in the evaluation as a condition of being in the programme. We will be clear that taking part in interviews and surveys is optional and that individuals can choose not to take part or to skip questions they do not wish to answer. For TAs, we will explain that observations are to learn how MTPB works rather than to assess their performance.

- **Incentives**: Due to the large ask of schools for this pilot evaluation (including baseline and endline pupil assessments), we have agreed with the EEF and the delivery team that a financial incentive of £500 for schools completing the IPE activities and pupil assessments will be offered.

- **Evaluation activities involving children**: We understand schools to be in *loco parentis* and able to consent to lesson observations and pupil assessments. We will provide materials for schools to inform parents about the programme and the evaluation. Parents will be informed via a parent letter that their child has been selected to take part in the programme aimed at children in reception who need extra support to develop a secure understanding of numbers and operations. They will also be informed that the programme will take place in normal school time and that they and their child won't need to do anything differently at home. Parents will be given the opportunity to opt out of their children's data being used for the evaluation. We will also provide separate parental consent letters for the child discussion groups, using opt-out consent. Due to the young age of the children concerned, we will verbally explain the purpose of our activities to children and let them know they can decide to not answer any questions or take part in any activities if they do not want to. Our team is experienced in carrying out research with young children and we will hold in-depth briefings on age-appropriate data collection. All researchers attending schools will have advanced DBS certification. The assessments will be carried out by qualified teachers experienced in working with this age group and with advanced DBS certification.

- **Risk of harm and safeguarding**: We do not expect the evaluation to increase the risk of harm beyond the low-level risks participants will face in delivering/participating in the pilot

evaluation. However, there is always the potential for participants to become upset or make a safeguarding disclosure. We will focus data collection directly on experiences of the programme and brief researchers on handling different situations that might arise. We will agree safeguarding and disclosure protocols with the EEF and the delivery team.

Data protection

Data collected for this evaluation will only be used for research purposes. We will store and handle data in line with General Data Protection Regulation (GDPR) and the Data Protection Act 2018. Only named individuals will have access to personal data and the team will comply with information security procedures that include preventative measures and processes for reporting, reviewing and responding to breaches. All data transferred to Oxford MeasurEd for the purposes of the evaluation will be done so via secure cloud software and Oxford MeasurEd will not transfer the data to any other party. We will securely delete personal data six months after the project ends.

We will pseudonymise all data, removing schools' and individuals' names prior to analysis. We will not report schools' or individuals' names in evaluation outputs, but some participants may be recognisable due to the small size of the pilot evaluation. We will communicate these limits to confidentiality and anonymity.

For this evaluation, Oxford MeasurEd is a data controller who also processes data. This means that we are responsible for deciding the purpose and legal basis for processing data. Our legal basis for data processing will be 'legitimate interests'. For any special categories of personal data, we will use the research exception. Data subjects will be the delivery team, TAs, teachers, children and SLT members. We will outline data protection procedures and safeguards and our legal bases for processing data during recruitment and will publish a privacy notice online and circulate it to all concerned parties.

After the evaluation ends, pupil data from the piloting of outcome measures will be matched with information from the Department for Education (DfE)'s National Pupil Database (NPD) and stored in the EEF archive, which is hosted in the Office for National Statistic (ONS)'s Secure Research Service (SRS). At this point, the EEF will become the data controller, and FFT Education (archive manger) the data processor, of the archived data. Anonymised data will be accessible to the wider research community for secondary analyses that provide public benefit and are in line with the missions of the EEF, DfE and ONS. Further matching to NPD and other administrative data may take place during this subsequent research. All evaluation data will be securely deleted from Oxford MeasurEd's systems no more than six months after submission of data to the EEF data archive and the final edited EEF report (December 2025).

Personnel

Delivery team

The programme will be delivered by a team from ELRS, based at Sheringham Nursery School:

Table 8 Delivery team

Name	Role	Institutional affiliation
Fliss James	Project delivery lead	East London Research School (ELRS)
Melissa Prendergast	Project delivery lead	East London Research School (ELRS)

Evaluation team

The evaluation will be delivered by a team from Oxford MeasurEd:

Table 9 Evaluation team

Name	Role	Institutional affiliation
Dr Lydia Marshall	Principal Investigator and Project Director	Oxford MeasurEd (Director of Research)
Dr Jonah Bury	Senior Evaluator and Project Manager	Oxford MeasurEd (Principal Consultant)
Robert Wishart	Impact Evaluation Expert	Oxford MeasurEd (Associate)
Sara Bashir Malik	Junior Evaluator and Project Coordinator	Oxford MeasurEd (Consultant)
Paulina Valenzuela	Junior Evaluator and Assessment Expert	Oxford MeasurEd (Consultant)

Risks

As part of our project management processes, the Project Director and Project Manager will develop and maintain a risk register, which will help us to anticipate and communicate risks to the EEF and the delivery team in a timely way and to set out mitigation strategies. The main risks to this project are detailed in Table 10.

Table 10 Initial risk register

Risk	Likelihood/impact	Mitigation/contingency
Evaluation delayed by delivery timeline slippages, including recruitment of schools to the programme	Likelihood: Medium Impact: Low	• Clear communication during inception outlining timelines and agreeing contingencies with the EEF and the delivery team • Flexibility to adapt timelines if needed
Evaluation delayed or disrupted by resourcing gaps	Likelihood: Low Impact: Low	• Use of established tools and processes to allocate, protect and adjust staff resource as needed • Wider pool of consultants with relevant skills and expertise to draw on if timelines or availability shift
Participant groups do not engage with the evaluation	Likelihood: Low Impact: High	• Emphasising value of pilot and findings to all participant groups • Minimising burden by keeping all data collection encounters short and offering flexible timings • Short and engaging communications with a clear ask

Schools are not willing to engage with outcome measure data collection (pupil assessments)	Likelihood: Medium Impact: High	• Desk-based review to consider acceptability of measures for use with young children and burden on schools as well as statistical properties • Liaison with the delivery team who felt external implementation of assessments would be more viable/acceptable to schools than school-implemented assessments due to current high workload for schools • Recruitment materials emphasising age-appropriateness of measures and processes • Financial incentive for schools • Qualitative research to ask school staff about attractiveness of possible incentives in a trial, including providing endline results to schools
We are unable to recruit a TA or teacher from every school to the qualitative interviews	Likelihood: Medium Impact: High	• Tailored recruitment emphasising the relevance and value of the research specifically for TAs and teachers • Liaising with schools so that recruitment materials can be sent during less demanding weeks of the school calendar • Contingency sampling strategy to be agreed with the EEF, e.g., interviewing teachers *and* TAs at a school rather than one or the other
Response bias in quantitative surveys	Likelihood: Medium Impact: Medium	• Recruitment materials emphasising that we want to hear from everyone about their experiences – good and bad • Invitations sent from Oxford MeasurEd to emphasise independence • Regular reminders using behavioural nudges to encourage all TAs/teachers to take part • Triangulation of findings with other data sources
Participants are unwilling to discuss barriers, challenges and negative effects	Likelihood: Medium Impact: Medium	• Recruitment materials emphasising that we want to hear from everyone about their experiences – good and bad • Scheduling by Oxford MeasurEd to emphasise independence • Use of individual interviews rather than paired or group interactions to encourage participants to speak openly • Appropriate reassurances about confidentiality • Using child-friendly techniques to encourage child engagement in group discussions
Analysis and reporting are of low quality	Likelihood: Low Impact: High	• Appropriately qualified team • Quality assurance (QA) by Director of Research, following established QA processes that can be agreed with the EEF • Pre-planning of QA by the EEF and peer review as per EEF protocols

Findings do not meet the needs of decision-makers (i.e., the EEF and ELRS stakeholders deciding whether to proceed to trial)	Likelihood: Low Impact: High	• Agreeing evaluation questions and design with the EEF and the delivery team at inception • Learning workshop to allow delivery team to input on emerging analysis and final reporting • EEF sign off on report structure and analysis plans

Timeline

Table 11 Timeline

Dates	Activity	Staff responsible/ leading
Feb 2023	IDEA workshop with delivery partners	EEF, Oxford MeasurEd
Feb – Mar 2023	Set-up meetings	EEF, Oxford MeasurEd, ELRS
Feb – Apr 2023	Desk-based review of programme documentation	Oxford MeasurEd
Feb – May 2023	Desk-based review of outcome measures	Oxford MeasurEd
Mar – Jul 2023	Study plan	Oxford MeasurEd
Apr – Jul 2023	Recruitment materials	Oxford MeasurEd, ELRS
Jun – Aug 2023	Ethical approval	Oxford MeasurEd
Jul – Dec 2023	Instrument development	Oxford MeasurEd
Sep – Oct 2023	Desk-based review of business as usual	Oxford MeasurEd
Oct – Dec 2023	School recruitment and teacher and TA enumeration	ELRS
Nov – Dec 2023	Pupil selection	Participating schools
Jan 2024	Pupil enumeration	Oxford MeasurEd
Jan 2024	Pupil baseline assessments	Oxford MeasurEd
Feb – Jun 2024	Training	ELRS
Feb – Jun 2024	Training observations	Oxford MeasurEd
Feb – May 2024	Collection of programmatic data	ELRS
Mar – May 2024	In-school delivery	Participating schools
Mar – May 2024	Delivery team staff interviews In-school observations School staff interviews Child discussion groups	Oxford MeasurEd
Jun – Jul 2024	Post-intervention surveys of TAs and teachers	Oxford MeasurEd
Jun – Jul 2024	Pupil endline assessment	Oxford MeasurEd
Jun – Jul 2024	Interim findings brief	Oxford MeasurEd

Jul 2024	Learning Workshop with delivery team	Oxford MeasurEd
Aug – Oct 2024	Final analysis of main evaluation data	Oxford MeasurEd
Sep 2024	Interim findings presentation for the EEF grant committee decision-making about proceeding to trial	Oxford MeasurED
Oct 2024	Longitudinal follow-up survey of TAs and teachers	Oxford MeasurEd
Nov 2024	Longitudinal follow-up survey analysis	Oxford MeasurEd
Dec 2024	Final draft report and presentation	Oxford MeasurEd

List of references

Baroody, A. J., & Lai, M. (2022). 'The development and assessment of counting-based cardinal number concepts'. *Educational Studies in Mathematics*, *111*(2), 185-205.

Brady, L. M., & Graham, B. (2019). *Social research with children and young people*. Bristol: Policy Press.

Casey, B., Erkut, S., Ceder, I., & Young, J.M. (2008). 'Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten'. *Journal of Applied Developmental Psychology,* 29(1), 29-48.

EEF (2018). *Improving mathematics in key stages two and three: Evidence review*. Available at: https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/maths-ks-2-3

EEF (2021a). *Improving mathematics in the Early Years and Key Stage 1*. Available at: https://d2tic4wvo1iusb.cloudfront.net/production/eef-guidance-reports/early-maths/EEF_Maths_EY_KS1_Guidance_Report.pdf?v=1694062549.

EEF (2021b). *Making best use of teaching assistants*. Available at: https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/teaching-assistants

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

EEF (2022b). *The ShREC approach – 4 evidence-informed strategies to promote high quality interactions with young children*. EEF blog. Available at: https://educationendowmentfoundation.org.uk/news/eef-blog-the-shrec-approach-4-evidence-informed-strategies-to-promote-high-quality-interactions-with-young-children

Hassinger-Das, B., Jordan, N. C., & Dyson, N. (2015). 'Reading stories to learn math: Mathematics vocabulary instruction for children with early numeracy difficulties'. *Elementary School Journal*, 116(2), 242-264.

Higgins, S., Katsipataki, M., Kokotsaki, D., Coleman, R., Major, L.E. and Coe, R. (2013). *The Sutton Trust- Education Endowment Foundation Teaching and Learning Toolkit*. Available at: https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit

Nicoletti, C., & Rabe, B. (2014). *School inputs and skills: Complementarity and self-productivity*. IZA Discussion Paper No. 8693.

Ofsted (2021). *Research review series: Mathematics*. Available at: https://www.gov.uk/government/publications/research-review-series-mathematics/research-review-series-mathematics

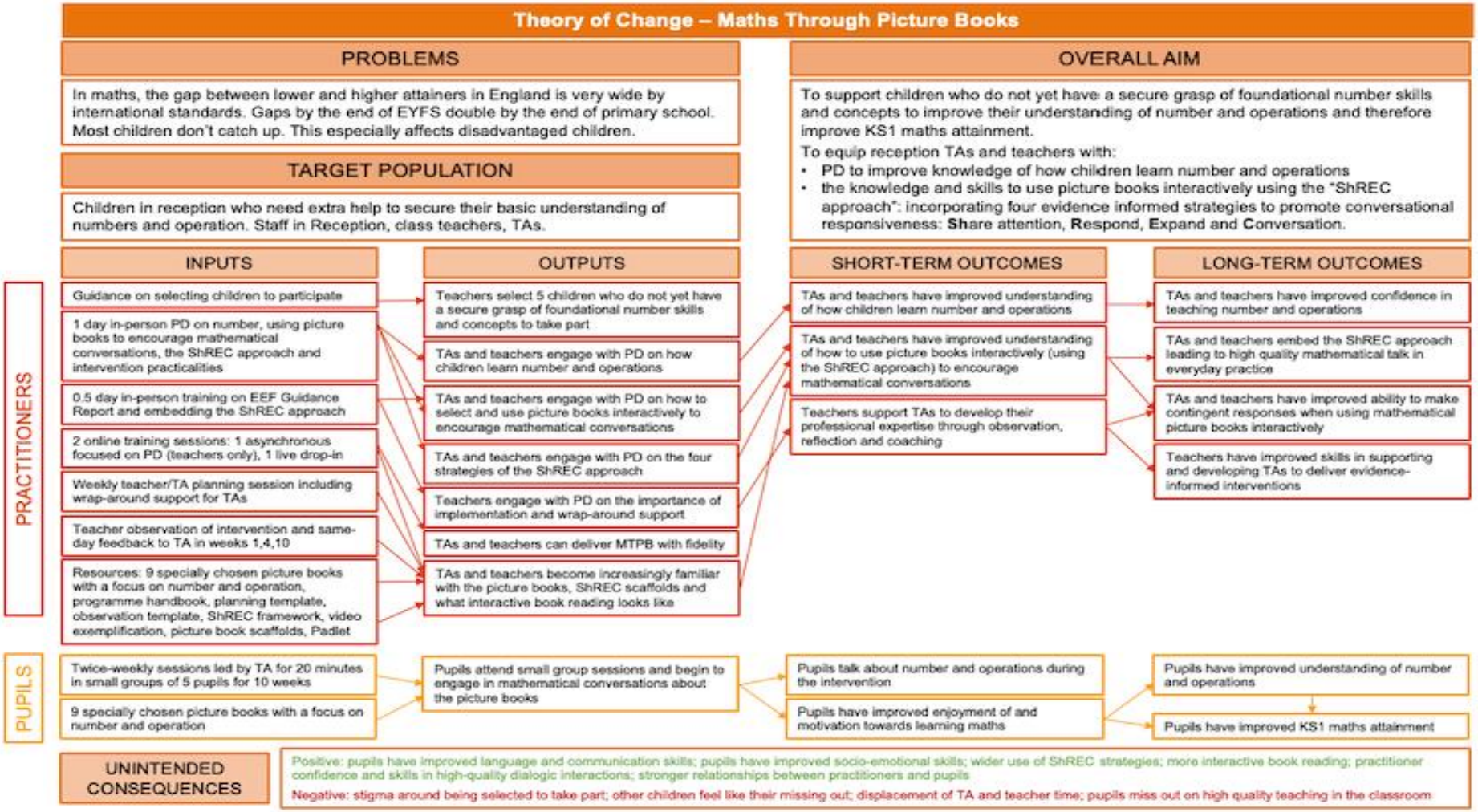
Purpura, D. J., Napoli, A. R., Wehrspann, E. A., & Gold, Z. S. (2017). 'Causal connections between mathematical language and mathematical knowledge: A dialogic reading Intervention'. *Journal of Research on Educational Effectiveness*, 10(1), 116-137.

Ritchie, J., Lewis, J., Nicholls, C. M., and Ormston, R. (eds.) (2013). *Qualitative Research Practice*. London: Sage.

Van den Heuvel-Panhuizen, M., Elia, I., & Robitzsch, A. (2016). 'Effects of reading picture books on kindergartners' mathematics performance'. *Educational Psychology*, 36(2), 323-346.

Appendix 1: Theory of Change

Appendix Figure 1 Theory of Change (developed by the delivery team with facilitation from the EEF and the evaluation team)



Appendix Figure 2 Assumed causal mechanisms (developed by the delivery team with facilitation from the EEF)

#	Assumption / Causal Mechanism	Where in the ToC does the assumption apply?		Evidence	Evidence Strength
1	Attendance at training will give practitioners a secure understanding of how children learn about number and operations	Output	Short-term outcome	Robust evidence: engaging in PD which includes a range of mechanimsms (building knowledge; motivating; developing teaching techniques; embedding practice) EEF PD GR; Robust evidence underpinning the Improving Mathematics in the EY & KS1 GR focused on importance of early mathematical development for children's current achievement, future learning and life success (Clements et al., 2013; Cross et al., 2009; Deans for Impact, 2019; Dooley et al., 2014). Evidence from practitioners involved in innovation pilot: reported in logs and focus group conversations.	Green
2	Attendance at training will give practitioners an improved understanding of effective teaching and learning strategies	Output	Short-term outcome	Weak-to-moderate evidence: The early years and KS1 evidence review of maths teaching identified 6 studies that examined the effect of using storybooks. Evidence from practitioners involved in innovation pilot: reported in logs and focus groups conversations.	Amber/Green
3	Improved motivation and confidence in teaching number and operations, overcoming 'maths anxiety'	Long-term outcome		Weak to moderate evidence: KS2/3 EEF Maths evidence review highlights the role of learners' attitudes and productive dispositions towards mathematics for successful learning. Dowker et al, 2016.	Amber/Red
4	Improved knowledge of how children learn about number and operations	Short-term outcome	Overall aim	Robust evidence: engaging in PD which includes a range of mechanimsms (building knowledge; motivating; developing teaching techniques; embedding practice) EEF PD GR: evidence underpinning the Improving Mathematics in the EY & KS1 GR.	Green
5	Use of specially chosen picture books leads to children showing improved motivation to learn about number. Children talk about number and operations during the intervention.	Short-term outcome	Long-term outcome	EEF Evidence review on Early Years & KS1 Mathematics Teaching: small but growing body of evidence to support use of storybooks to teach mathematics. 6 studies all involved the use of specially chosen or specially designed storybooks. van den Heuvel-Panhuizen & Elia (2012) developed a framework for evaluating the suitability of picture books for children's mathematical development, evidence-informed guidance is available from the US-based Development and Research in Early Math Education (DREME) website: https://dreme.stanford.edu/. Evidence from practitioners involved in the innovation pilot: reported children demonstrated high levels of motivation and increased talk about number and operations.	Amber/Green
6	Engagement in the programme leads to improved confidence in teaching number and operations in TAs	Long-term outcome		EEF evidence review on effective PD. The programme is developed using the balanced approach to ensure mechanisms from all four categories are incorporated. Behaviour change research indicates sustained changes in practice can be difficult to achieve.	Amber/Green
7	Engagement in the intervention leads to effective practices being embedded and sustained in classroom practice.	Long-term outcome		Evidence as above. Behaviour change research indicates sustained changes in practice can be difficult to achieve.	Amber/Green
8	Children are better prepared for KS1.	Long-term outcome		Robust evidence underpinning the Improving Mathematics in the EY & KS1 GR focused on importance of early mathematical development for children's current achievement, future learning and life success (Clements et al., 2013; Cross et al., 2009; Deans for Impact, 2019; Dooley et al., 2014). Evidence from practitioners involved in the pilot: pracitioners reported children demonstrated high levels of motivation and increased talk about number and operations.	Green
9	Children have improved understanding of number and operations	Long-term outcome		Robust evidence underpinning the Improving Mathematics in the EY & KS1 GR focused on importance of early mathematical development for children's current achievement, future learning and life success (Clements et al., 2013; Cross et al., 2009; Deans for Impact, 2019; Dooley et al., 2014). Evidence from practitioners involved in the pilot: practitioners reported children demonstrated high levels of motivation and increased talk about number and operations.	Green

Key: **green** – the evidence base is very strong, **green/amber** – the evidence base is strong, **red/amber** – the evidence base is developing, **red**- the evidence base is limited.

Appendix Figure 3 Contextual Assumptions (developed by the delivery team with facilitation from the EEF)

#	Contextual Assumption	Assumption Strength	Assumption Risk
1	Schools read the requirements for participation in the intervention and sign-up fully knowing what is expected of them	Amber/Green	Red
2	Teachers and teaching assistants can access the twilight online training and can attend all of it; they can attend the one and a half day's f2f training which will require cover	Green	Red
3	Teachers and TAs have gaps in their knowledge or hold misconceptions about the most effective ways to support children to learn about numbers and operations	Green	Amber/Red
4	Delivery team can source enough copies of specially chosen books	Amber/Green	Red
5	Teachers and TAs are motivated to enage in the training	Amber/Green	Amber/Red
6	Teachers and TAs have time to deliver the sessions as intended	Amber/Green	Red
7	Teachers are able to monitor consistent implementation, model strategies and offer feedback to TAs, supporting reflection and planning	Amber/Green	Red
8	Appropriate children are selected for the intervention (children struggling with number - not children likely to achieve the Number ELG)	Green	Amber/Green
9	Children attend the intervention sessions	Amber/Green	Amber/Red
10	Losing the TA to run the intervention means that time for other important learning is lost, like early reading	Amber/Green	Amber/Red
11	Schools can find a quiet space and timetable slot to deliver the sessions.	Amber/Green	Amber/Red
12	SLT are supportive of the intervention.	Amber/Green	Amber/Red
13	Staff illness impacts intervention delivery	Green	Amber/Red
14	School have the staffing structure and capacity to commit fully to the programme	Amber/Green	Amber/Red
15	Schools who start the programme remain committed and engaged	Amber/Green	Amber/Red
16	Teachers and TAs understand the intervention and what they are being asked to deliver	Green	Amber/Red
17	Teachers and TAs have a positive professional relationship and coaching supports behaviour change	Amber/Green	Amber/Red

Assumption strength key: **green** – this assumption will hold in the vast majority of circumstances where the programme is delivered, **green/amber** – this assumption will hold in most of the circumstances where the programme is delivered, **red/amber** – this assumption will often not hold in the circumstances where the programme is delivered, **red** – there is a good chance of this assumption not holding / do not know whether this assumption will hold or not.
Assumption risk key (if this assumption does not hold…): **green** – the programme could continue to be delivered with very minor impact, **green/amber** – the programme could continue to be delivered, but the impact would be substantial, **red/amber** – the programme could continue to be delivered, but without fidelity to original design; **red** – the programme could not be delivered.

Appendix 2: Review of Outcome Measures

This note outlines the findings of a review of the availability and suitability of different numeracy outcome measures for the MTPB pilot evaluation.

Identification of outcome measures

In the first stage we identified potential numeracy outcome measures that are designed for use with reception age pupils (four to five years old). The measures were identified from multiple sources:

- Education Endowment Foundation (EEF) literature and guidance, such as the review of assessment measures in the early years and early years measures database.
- Commercial assessment providers such as GL Assessment and Hodder Education.
- Outcome measures used in previous evaluations of early years numeracy.
- Existing national assessments: the only assessment conducted with this age group is the Reception Baseline Assessment (RBA). However, this data will not be available to schools or to researchers in the National Pupil Database (NPD) at pupil level and was therefore excluded.

In total, 22 outcome measures were identified for the long list.

Assessing suitability and quality

Suitability was determined by whether the assessment has been developed for use with pupils in this age group, that it covers the domains of numeracy (understanding of number and operations) that the MTPB programme is aiming to change and whether it has been designed for use in the UK. It also included logistical considerations such as how the test is administered (for example, oral or paper administration), and who can administer the test (such as school staff or an external assessor).

Quality was determined by the sample the measure was previously developed/tested with, validity and reliability.

Use of tool was assessed by exploring what the measure had previously been used for (i.e., ongoing assessment, to inform teacher practice, national benchmarking or evaluation).

Suitability, quality and use was screened using a bespoke screening tool developed for the outcome scoping exercise.

Challenges

There are several challenges with identifying an appropriate outcome for this age group. The key challenges that were encountered were:

- Finding a measure that is appropriate for use with 4-year-olds, which is particularly problematic for baseline measurement.

- Finding measures that were suitable for use in the UK, as many measures were developed in the US.
- Finding detailed information about the specific domains of numeracy that are captured by different measures as very limited information is published.

A key consideration for the evaluation is therefore whether we can use the same outcome measure at baseline and endline. We therefore consider options for baseline and endline measurement separately.

Shortlisted baseline measures

The shortlisted baseline measures and their relative strengths and limitations are outlined in Appendix Table 1.

Appendix Table 1 Shortlisted baseline outcome measures

Measure	Strengths	Limitations
New Progress in Understanding Mathematics Assessment (PUMA)	• 4 – 11 years • Has been designed for use in the UK • UK norms tested • Separate tests available for autumn, spring, summer terms • Can be administered digitally • 30 – 40 minutes to assess whole class • £17.50 per 10-copy pack of termly tests • Tested on a sample size of 10,000 in 2019/20[4] • Excellent construct validity	• Domains covered unclear • Administered as a group but individual reports available • Information on reliability not available
Pearson Wechsler Individual Achievement Test - Third UK Edition (WIAT-III UK)	• Suitable for 4 – 25 years 11 months • Has been designed for use in the UK • UK norms tested • Covers required domains of numeracy and basic operations • Can be administered digitally • 10 minutes to administer each subtest • Tested on a sample size of 3,000[5] • Moderate – strong construct validity • Good – excellent internal consistency	• £783.55 for 25 tests – discount may be possible for large scale assessment (i.e., efficacy trial) • Unclear if only one subtest can be bought and administered (clarification sought) • Only the Math Problem Solving subtest is suitable for 4 – 17+ years, Numeracy subtest is suitable for 5 – 17+ years, Maths fluency for 6 – 17+ years

Shortlisted endline measures

The shortlisted endline measures and their relative strengths and limitations are outlined in Appendix Table 2.

[4] More information about New PUMA can be found on the publisher's website here.
[5] More information about WIAT-III UK can be found on the EEF website here.

Appendix Table 2 Shortlisted endline outcome measures

Measure	Strengths	Limitations
PUMA	• 4 – 11 years • Has been designed for use in the UK • UK norms tested • Separate tests available for autumn, spring, summer terms • Can be administered digitally • 30 – 40 minutes to assess whole class • £17.50 per 10-copy pack of termly tests • Tested on a sample size of 10,000 in 2019/20 • Excellent construct validity	• Domains covered unclear • Administered as a group but individual reports available • Information on reliability not available
WIAT-III UK	• Suitable for 4 – 25 years 11 months • Has been designed for use in the UK • UK norms tested • Covers required domains of numeracy and basic operations • Can be administered digitally • 10 minutes to administer each subtest • Tested on a sample size of 3,000 • Moderate – strong construct validity • Good – excellent internal consistency	• £783.55 for 25 tests – discount may be possible for large scale assessment (i.e., efficacy trial) • Unclear if only one subtest can be bought and administered (clarification sought) • Only the Math Problem Solving subtest is suitable for 4 – 17+ years, Numeracy subtest is suitable for 5 – 17+ years, Maths fluency for 6 – 17+ years
Wide Range Achievement Test – Fifth Edition (WRAT5)	• UK norms tested • Covers required domains of numeracy and basic operations • Can be administered digitally • Approximately 15 – 25 minutes for ages 5 – 7 • Tested on a sample size of 2,000[6] • Excellent construct validity • Average – excellent reliability • $431 for 25 tests, possible to buy only Maths subtest (price unclear)	• 5 – 85+ years

Recommendations

Based on the currently available information, the **WIAT-III UK** assessment is deemed most suitable for the MTPB evaluation.

There are three subtests related to mathematics, which are:

- **Math Problem Solving** (4 – 17+ years) - Measures untimed maths problem-solving skills in the following domains: basic concepts, everyday applications, geometry, and algebra. The student provides oral and pointing responses.
- **Numeracy** (5 – 17+ years) - Measures untimed, written maths calculation skills in the following domains: basic skills, basic operations with integers, geometry, algebra, and calculus.

[6] Pearson provide a more comprehensive overview of WRAT5 UK, which can be accessed here.

- **Math Fluency** (6 – 17+ years) - Measures the speed and accuracy of a student's maths (addition) calculations. The student solves written addition problems within a 60-second time limit.

At baseline, only the math problem solving subtest is appropriate given the age of the cohort being tested. At endline, both the math problem solving subtest and numeracy subtest could be used. It may be beneficial to use both subtests as it covers a greater range of mathematics skills and may reduce the likelihood of encountering ceiling effects.

The WIAT-III UK maths problem solving subtest uses a picture book, where the assessor asks a question related to the image on the page. The assessment has a large number of items but uses a stop rule; when three or more questions are answered incorrectly, the test ends. This means it is unlikely to suffer from ceiling effects. It covers basic understandings of shape and number, with limited application of operations. The numeracy subtest provides further assessment of understanding of operations using a response booklet that the child will complete.

Contrastingly, the WRAT5 uses a response booklet, and the questions are more similar to the numeracy subtest of WIAT-III UK. Based on demo materials shared by Pearson Education, it is not suitable for use with the sample at baseline for this reason.

Consequently, the preferred approach would be to use the WIAT-III UK maths problem solving test at baseline, and both the WIAT-III UK maths problem solving and numeracy tests at endline.

The tests can be administered digitally or with a paper copy. A consideration is the financial cost of the test, as it is relatively expensive compared with other measures.

Appendix 3: Child Selection Guidance and Observation Grid

Child selection guidance

To help you to select the **5** children that would benefit most from taking part in the Maths Through Picture Books pilot we would like you to carry out a task with **all** children in your reception class.

This is a short and engaging task, based on current research that identifies two key important indicators which can predict later success in mathematics: *counting out a specific amount from a larger group of objects and understanding number symbols.*

The aim of the task is for the child to:

- *count out a set of a given number;*
- *match the correct corresponding numeral.*

What you will need:

- Teddy bear or other soft toy/puppet
- Pennies (or other objects such as buttons or jewels) in a pile
- Wooden/plastic numerals 1-9
- Observation grid to record the child's responses.

Task instructions:

Each child will be asked to engage in the following task which involves giving a teddy bear (or other soft toy or puppet) a specific number of pennies from a larger amount:

"Can you give Mr Bear 3 pennies?"

Observe what the child does and record the child's response on the observation grid provided. Circle the number if they gave the correct amount. Briefly note how the child engages in the activity for example are they confident to join in?

Once the child has given you the pennies, ask them to find the right numeral to go with that amount:

"Mr Bear has 3 pennies...I wonder if you can help Mr Bear find the right number to go with his pennies?"

Record the child's response in the second column of the grid. Circle the numeral if they matched it correctly.

It will be helpful if you note down anything the child says or any particular strategies they use when engaging in the task.

If you identify more than five children who find the tasks challenging, please use your professional judgement and knowledge of each individual child to make your decision as to which children would benefit most from participating in the pilot.

Observation grid

Child's name	**Give number**	**Match numeral**	**Notes** ● *Circle which numbers they give successfully.* ● *Circle which numerals they match correctly.* ● *Pay attention to the child's attitude and confidence.*
Example: Adam	3 5 9	3 5 9	Happy and confident to join in with the activity. Counted 3 pennies and found numeral 3. Counted to 5 but produced 4. Identified the numeral 4. When asked to count out 9 he grabbed a handful of pennies commenting "I got lots!" When counting beyond 5, skipped some numbers: '5,6,8,10.'

Appendix 4: Outcome analysis plan

We will analyse quantitative data from the outcome measure piloting to provide key statistical information.

Primary outcome analysis

The primary outcome is maths attainment, measured by the WIAT-III total raw score, which has a scale of 0-133. As is outlined above, to estimate the variance explained by the baseline of the primary outcome, we will estimate a multi-level linear regression model. The regression formula for this model will be:

$$Y_{ijt_1} = \beta_0 + \beta_1 Y_{ijt_0} + \mu_j + \varepsilon_{ij}$$

Where:

- The dependent variable *Y* is Maths attainment at endline (t_1) for individual *i* in school *j*.
- β_0 represents the regression intercept.
- β_1 represents the regression coefficient for Maths attainment at baseline (t_0).
- μ_j represents the random intercept for schools.
- ε_{ij} is the random error term.

The ICC formula is outlined in the relevant section later in the analysis plan.

Secondary outcome analysis

The secondary outcomes are mathematical problem solving and numeracy, measured by the problem solving and numeracy subscales of WIAT-III UK. The analysis will mirror the primary outcome analysis, by estimating the parameters outlined above for each subscale.

Missing data

Individual item data is unlikely to be missing because of the way the test is administered. A stop-rule is used, so that the test is stopped when a child answers four consecutive items incorrectly. The remaining items are therefore considered to have a value of zero. The most likely cause of missing data will therefore arise from pupil absence, where they are unavailable on the day of testing or on any of the "mop up" testing days. Nevertheless, the overall proportion of missing cases will be reported, and any patterns in missing data at item level will be assessed.

As the pilot does not intend to estimate impacts, no imputation is necessary in the presence of missing data.

Compliance

The pilot is not estimating impact, and therefore, no complier average causal effect (CACE) will be estimated. Nevertheless, we will consider the data collected from the pilot concerning training attendance and other measures of fidelity to consider how compliance could be assessed in an efficacy trial.

Intra-cluster correlations (ICCs)

The ICC will be estimated using the postestimation command estat icc. This uses the following formula.

$$\rho = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} = \frac{\sigma_B^2}{\sigma_T^2}$$

Where:

- σ_B^2 is the between school variance.
- σ_W^2 is the within school variance.
- σ_T^2 is the total variance.

Appendix 5: Survey analysis plan

The survey sample includes a census of all teachers and TAs participating in the pilot for surveys at two time points: post-intervention in June 2024 and longitudinal follow-up in October 2024.

Descriptive analysis

We will analyse quantitative survey descriptively, using Stata, to report on the range of experiences of the pilot and perceived outcomes and explore whether these differ between key participant groups. Data will be analysed using unweighted frequencies and averages.

We will also use cross-tabulation analysis to test relationships between experiences and outcomes – for example whether perceived outcomes differed between groups experiencing different barriers to implementation.

Longitudinal analysis

We will link responses from the longitudinal follow-up survey to the post-intervention survey using unique IDs, in order to explore sustained practice and the relationships between outcomes over time.

We will report on missing data at both time points and on attrition between the post-intervention and longitudinal follow-up surveys and explore whether missingness or attrition was equally distributed by background characteristics. Missing data will not be imputed.

Detailed analysis plan

Appendix 5 Table 12 sets out the analyses we plan to conduct against each research question. The surveys can be found in Appendix 6: TA post-intervention survey and Appendix 7: Teacher post-intervention survey.

Output from all analyses will be included in an annex to the evaluation report.

Appendix 5 Table 12 Analysis plan

Variables		Research questions	Analysis	Survey questions

Evidence of promise	**Outcomes:**
	• Practitioners' understanding
	• Practitioners' confidence
	• Children's enjoyment
	• Children's motivation
	• Unintended consequences
		1. Is there an improved understanding among TAs and Teachers, of how children learn about numbers and operations?
		2. Is there an increased confidence in teaching maths among TAs and Teachers?
		3. Do children talk about number and demonstrate increased enjoyment of and motivation towards learning mathematics?
		5. Are there any unintended consequences of participation in this programme?
			• Overall frequencies post-intervention and at longitudinal follow-up
			• Averages post-intervention and at longitudinal follow-up
			• Sub-group comparisons (chi-sq)
			  ○ Role (teacher vs TA)
			  ○ Years of experience
			  ○ Highest level of qualification
			  ○ Age
			  ○ Ethnicity
			• Trajectory of reported outcomes post-intervention to reported outcomes at longitudinal follow-up (overall and sub-group comparisons)
			• Chi-square testing to explore relationship between outcomes and barrier and facilitator variables
				TA post-intervention: Q26-36
				Teacher post-intervention: Q29-41
				TA longitudinal follow-up:
				Teacher longitudinal follow-up:

	Intention to sustain	4. Is teaching practice embedded and sustained after 10 weeks?	• Overall frequencies post-intervention and at longitudinal follow-up
			• Sub-group comparisons (chi-sq)
			  ○ Role (teacher vs TA)
			  ○ Years of experience
			  ○ Highest level of qualification
			  ○ Age
			  ○ Ethnicity
			• Trajectory of intention to sustain post-intervention to actual sustained practice at longitudinal follow-up (overall and sub-group comparisons)
			• Chi-square testing to explore relationship between intention to sustain or sustained practice and self-reported practitioner understanding and confidence
				TA post-intervention: Q37
				Teacher post-intervention: Q42
				TA longitudinal follow-up:
				Teacher longitudinal follow-up:

	Difference to business as usual	6. How different is the programme to business as usual?	• Overall frequencies post-intervention • Averages post-intervention	**TA post-intervention:** Q41-43 **Teacher post-intervention:** Q43-45
Fe asi	**Fidelity**	7. Is the programme delivered as intended?	• Overall frequencies post-intervention	**TA post-intervention:** Q3, Q9, Q14-17, Q20-21

		9. What adaptations are participants making to the programme and why?	• Averages post-intervention	**Teacher post-intervention:** Q7, Q14, Q18-21, Q24-25

	Facilitators Challenges:
	• barriers
	• financial costs
	• time spent on intervention
		8. What contextual factors affect implementation?
		12. Is the programme affordable for schools?
			• Overall frequencies post-intervention
			• Averages post-intervention
			• Sub-group comparisons (chi-sq)
			  • Role (teacher vs TA)
			  • Years of experience
			  • Highest level of qualification
			  • Age
			  • Ethnicity
				TA post-intervention: Q7-8, Q10, Q12-13, Q18-19, Q22, Q24-25, Q38-40
				Teacher post-intervention: Q12-13, Q15-17, Q22-23, Q26-28, Q46-48

	Pupil selection	10. How appropriate are the pupil selection/eligibility criteria?	• Overall frequencies post-intervention • Averages post-intervention	**Teacher post-intervention:** Q3-6

	Usefulness of support	11. Are programme inputs (training sessions, physical resources, additional support) accessible and useful to practitioners?	• Overall frequencies post-intervention
			• Averages post-intervention
			• Sub-group comparisons (chi-sq)
			  • Role (teacher vs TA)
			  • Years of experience
			  • Highest level of qualification
			  • Age
			  • Ethnicity
				TA post-intervention: Q4-6, Q11, Q23
				Teacher post-intervention: Q8-11

Subgroup analysis

As indicated in Appendix 5 Table 12, we plan to perform subgroup analyses using chi-square tests to test the hypothesis that the outcomes, experiences, and perceptions of the programme may differ between practitioner groups. We plan to conduct subgroup analyses based on years of experience and highest level of qualification to account for potential differences in participants' backgrounds and expertise, which may influence their experiences and perspectives. Additionally, we will explore variations in age and ethnicity, based on the power dynamics and social and cultural factors that could shape interactions with the intervention – particularly the training and teacher-TA support. By stratifying our data into subgroups, we will investigate potential associations and uncover any statistically significant differences between groups. Unique IDs will be used to match the survey data to the background information collected from participating teachers and TAs at the inception of the evaluation. All 19 teachers and 16 TAs agreed to provide this information.

The distribution of the sample for teachers and TAs is given in Appendix 5 Table 13, using that background data.

Appendix 5 Table 13 Distribution of sample

		Teachers	TAs	Total
Years of experience	Less than 1 year		5	5
	1-2 years	4	3	7
	3-4 years	2	3	5
	5-10 years	4	3	7
	More than 10 years	9	2	11
	Missing	-	3	3
Highest level of qualification	GCSE or equivalent	-	1	1
	A-level or equivalent	-	1	1
	Certificate of higher education	-	4	4
	Diploma of higher education	-	4	4
	Bachelor's degree	17	3	20
	Master's degree	2	1	3
	Other	-	2	2
	Missing	-	3	3
Gender	Male	2	0	2
	Female	17	15	32
	N/A	-	4	4
Age	18-30	8	3	11
	31-40	5	4	9
	41-50	2	7	9
	51-60	4	2	6
	Missing	-	3	3
Ethnicity	White	15	9	24
	Asian or Asian British	1	4	5
	Black, Black British, Caribbean or African	3	2	5
	Arab	-	1	1
	Missing	-	3	3

Since there are no male TAs and only two male teachers, subgroup analysis will not include gender. Subgroup analysis will also be undertaken to explore differences between the views and experiences of teachers and TAs.

Given the small sample size and categories, we will only carry out sub-group analysis for the overall sample, not for teachers and TAs separately, unless a category of questions only applies to either teachers or TAs.

Smaller categories (n≤5) within the data will be combined, where possible, to protect the identity of participants when reporting information. Pre-empting the possibility of small cell counts in a cross-tab analysis based on the small categories identified Appendix 5 Table 13, we plan to combine the categories as follows:

- Years of experience: We will categorise this as 4 years or less (n=17) and more than 4 years of experience (n=18).
- Highest level of qualification: We will categorise this as educated to a degree level or higher (n=23) and not educated to a degree level (n=12).
- Age: We will categorise this as 18-40 (n=20) and 41-60 (n=15).
- Ethnicity: We will categorise ethnicity as White (n=24) and non-White (n=11).

In addition to the categories identified above, we will also further supress categories where similar responses for small groups might make their responses identifiable to maintain anonymity.

All sub-group analyses should be treated with caution due to the small sample size. In particular, we will note that findings may be skewed given that most teachers are White and all are educated to a degree level or higher.

Appendix 6: TA post-intervention survey

MATHS THROUGH PICTURE BOOKS

Follow-up survey

Welcome to our follow-up survey to the Maths Through Picture Books (MTPB) programme. We are inviting all Teaching Assistants and teachers who have taken part in the MTPB pilot to complete this survey. The survey should take around **15 minutes** to complete.

This is the first of two surveys that will be used to help us understand how well the MTPB programme has worked in schools. We will email to invite you to complete the second survey in October 2024, which will be approximately 5 minutes long.

There are no right or wrong answers – we are interested in how the programme works and how useful it has been. Please answer as honestly as you can. Your responses will only be visible to the evaluation team at Oxford MeasurEd.

You can find out more about our evaluation and the surveys in the email we sent you inviting you to take part in the survey. You can find out more about how we will use your data in the privacy notice.

If you have any questions or concerns about the evaluation, please contact the Project Director, Lydia Marshall, at mailto:MTPB@oxfordmeasured.co.uk.

Q1 Please enter the unique ID assigned to you. You can find this in the email which contains the link to this survey.

[ ]

Q2 Are you happy to continue?
Yes >> Q3
No

Thank you – opt out

Thank you for your interest in the evaluation. If you change your mind about taking part in the survey, you can come back using the same link to complete the survey.

TRAINING AND SUPPORT

In this section, we want to know about how you found the training and support you received to deliver the MTPB programme at your setting.

Q3. What **training/support** did you access? Select all that apply.

- Full-day in-person professional development on 1st March 2024
- Online drop-in support session on 21st March 2024
- Half-day in-person professional development on 14th June 2024

- None of these

Q4. How **useful** would you say the training and support from the MTPB delivery team (Fliss James and Melissa Prendergast) has been?

- Not useful at all
- Somewhat useful
- Useful
- Very useful

Q5. How **confident** would you say you felt about **planning the MTPB classroom sessions** after the initial training and support you received from the MTPB delivery team?

- Not at all
- Somewhat
- Quite a bit
- A lot

Q6. How much would you say you **understood** about **how you should deliver the MTPB classroom sessions** after the initial training and support you received from the MTPB delivery team?

- Not at all
- Somewhat
- Quite a bit
- A lot

Q7. Did anything make it **harder** for you to **access the training/support** from the MTPB delivery team? Select all that apply.

- Lack of support from the MTPB delivery team
- Lack of support from your Senior Leadership Team (SLT)
- Lack of resources
- Lack of time
- Other commitments/priorities
- Other, please specify

Q8. Did anything make it **easier** for you to **access the training/support** from the MTPB delivery team? Select all that apply.

- Support from the MTPB delivery team
- Support from your Senior Leadership Team
- Resources (e.g., programme handbook, video exemplifications)
- Other, please specify

PLANNING

Next, we want to know about your experiences while planning the MTPB classroom sessions.

Q9. How often did you have the **weekly planning sessions** with your teacher throughout the 10-week period of the pilot?

- Every week

- At least every second week
- Fewer than every second week
- Only once or twice
- Never

Q10. How much time did you spend on **weekly planning sessions** on average?

- Less than 15 minutes
- 15 – 30 minutes
- 31 – 60 minutes
- More than 1 hour

Q11. How **useful** would you say was the support you received from the teacher in planning sessions?

- Not useful at all
- Somewhat useful
- Useful
- Very useful

Q12. Did anything make it **harder** to have the **weekly planning sessions**? Select all that apply.

- Lack of support from teacher
- Lack of support from Senior Leadership Team (SLT)
- Lack of resources
- Lack of time
- Other commitments/priorities
- Other, please specify

Q13. Did anything make it **easier** to have the **weekly planning sessions**? Select all that apply.

- Support from teacher
- Support from Senior Leadership Team (SLT)
- Resources (e.g., programme handbook, video exemplifications)
- Other

DELIVERY

Now, we want to learn about what the classroom delivery of MTPB looked like in your school. We want to know what worked in practice – so please answer honestly.

Q14. How often did **two** MTPB classroom sessions take place in a week throughout the 10-week period of the pilot?

- Every week
- At least every second week
- Fewer than every second week
- Only once or twice
- Never

Q15. Did the MTPB classroom sessions last for **20 minutes** each?

- No, they were less than 20 minutes long
- No, they were more than 20 minutes long

- No, they varied (sometimes less than and sometimes more than 20 minutes long)
- Yes, they were always 20 minutes long

Q16. Did the sessions run in **a dedicated, quiet space** away from the classroom?

- No, never
- Yes, sometimes
- Yes, always

Q17. Did you deliver the MTPB classroom session using the **Share attention, Respond, Expand, Conversation (ShREC) framework**?

- No, never
- Yes, sometimes
- Yes, always

Q18. Do you think anything made it **harder** for you to **deliver** the MTPB classroom sessions? Select all that apply.

- Poor child selection
- Lack of support from teacher
- Lack of support from Senior Leadership Team (SLT)
- Lack of resources
- Lack of time
- Other commitments/priorities
- Other, please specify

Q19. Do you think anything made it **easier** for you to **deliver** the MTPB classroom sessions? Select all that apply.

- Support from your teacher
- Support from your Senior Leadership Team (SLT)
- Resources (e.g., programme handbook, video exemplifications)
- Other, please specify

IN-CLASS FEEDBACK AND COACHING

In this section, we would like to learn more about how the teacher-led observation and Teacher-TA feedback and reflection sessions took place.

Q20. Did you have in-class feedback and coaching sessions with the teacher?

- Yes >> Q21
- No >> Q26
- Sometimes >> Q21

Q21. Did the **in-class feedback and coaching sessions** with the teacher take place on the **same day** as the observation by teacher?

- Yes
- No
- Sometimes

Q22. How much time did you spend on the **in-class feedback and coaching session** with the teacher on average?

- Less than 15 minutes
- 15 – 30 minutes
- 31 – 60 minutes
- More than 1 hour

Q23. How **useful** would you say were the **in-class feedback and coaching sessions** from the teacher?

- Not useful at all
- Somewhat useful
- Useful
- Very useful

Q24. Did anything make it **harder** to carry out the **in-class feedback and coaching** sessions? Select all that apply.

- Lack of support from your Senior Leadership Team (SLT)
- Lack of resources
- Lack of time
- Other commitments/priorities
- Other, please specify

Q25. Did anything make it **easier** to carry out the **in-class feedback and coaching** sessions? Select all that apply.

- Support from your Senior Leadership Team (SLT)
- Resources (e.g., programme handbook, video exemplifications)
- Other

OUTCOMES FOR TAs

Now, we want to ask you about any potential benefits and consequences for you as a TA as a result of taking part in the MTPB programme.

Q26. Has your own **understanding** of how children learn about numbers and operations improved as a result of taking part in the MTPB programme?

- No, not at all
- Yes, somewhat
- Yes, quite a bit
- Yes, a lot

Q27. Has your own **confidence** in teaching maths improved as a result of taking part in the MTPB programme?

- No, not at all
- Yes, somewhat
- Yes, quite a bit
- Yes, a lot

Q28. Were there any **other benefits** for you as a TA as a result of taking part in the MTPB programme? Please specify.

Q29. Have you noticed any change in your **relationships** with children in your class as a result of taking part in the MTPB programme?

- No, not at all
- Yes, they have got a bit worse
- Yes, they have got a lot worse
- Yes, they have got a bit better
- Yes, they have got a lot better

Q30. Were there any **negative consequences** for you as a TA as a result of taking part in the MTPB programme? Please specify.

OUTCOMES FOR CHILDREN

Next, we want to ask you about the potential benefits and consequences for children as a result of taking part in the MTPB programme.

Q31. For how many of the participating children would you say their **enjoyment** of maths has increased as a result of taking part in the MTPB programme?

- None
- 1 or 2
- 3 or 4
- All of them

Q32. For how many of the participating children would you say their **motivation** towards learning maths has increased as a result of taking part in the MTPB programme?

- None
- 1 or 2
- 3 or 4
- All of them

Q33. Were there any **other benefits** for participating children as a result of taking part in the MTPB programme? Select all that apply.

- Improved social emotional learning

- Improved language and communication
- Improved child-adult relationships
- Other, please specify

Q34. Were there any **negative consequences** for participating children as a result of taking part in the MTPB programme? Select all that apply.

- Time away from other activities
- Stigma of being selected
- Feeling left out for not being selected
- Reduced enjoyment of maths
- Reduced motivation towards learning maths
- Other, please specify

Q35. Do you think any groups of participating children **benefitted more** from the MTPB programme than other participating children? Select all that apply.

- Children with English as an Additional Language (EAL) benefitted more than others
- Children with Special Educational Needs (SEN) benefitted more than others
- Children entitled to Free School Meals (FSM) benefitted more than others
- Other group(s) benefitted more than others, please specify

Q36. Do you think any groups of participating children **benefitted less** from the MTPB programme than other participating children? Select all that apply.

- Children with English as an Additional Language (EAL) benefitted less than others
- Children with Special Educational Needs (SEN) benefitted less than others
- Children entitled to Free School Meals (FSM) benefitted less than others
- Other, please specify

INTENTION TO SUSTAIN

We are interested in whether you will continue using some of the practices you have learned from the MTPB programme.

Q37. Do you intend to **continue** using any of the following practices now that the MTPB programme has ended? Select all that apply.

- Use of picture books to teach maths
- Use of ShREC technique
- Small-group reading of maths picture books to support children falling behind

COSTS

We are now interested in finding out if you, as an individual, incurred any costs as a result of taking part in the programme. Please do **NOT** include the costs incurred by the school, e.g., the cost of a cover while you attended training.

Q38. Did you face any costs for taking part in and delivering the programme?

- Yes >> Q39
- No >> Q41

Q39. Please list the costs you incurred.

Q40. How much did you spend in total?

SCHOOL INFORMATION

Finally, we want to ask some information about how maths is normally taught at your school.

Q41. How **different** is the MTPB programme to how maths is taught normally to reception-age children?

- Very different
- Somewhat different
- Somewhat similar
- Very similar

Q42. Why?

Q43. Is your school currently involved in any **other** early years maths programmes?

- Yes, please specify
- No
- Don't know

Thank you

Thank you for completing this survey. Your answers will be really valuable to the evaluation and we appreciate your time.

If you have any questions, please contact the Project Director, Lydia Marshall, at mailto:MTPB@oxfordmeasured.co.uk.

Appendix 7: Teacher post-intervention survey

MATHS THROUGH PICTURE BOOKS

Follow-up survey

Welcome to our follow-up survey to the Maths Through Picture Books (MTPB) programme. We are inviting all Teaching Assistants (TAs) and teachers who have taken part in the MTPB pilot to complete this survey. The survey should take around **15-20 minutes** to complete.

This is the first of two surveys that will be used to help us understand how well the MTPB programme has worked in schools. We will email to invite you to complete the second survey in October 2024, which will be approximately 5 minutes long.

There are no right or wrong answers – we are interested in how the programme works and how useful it has been. Please answer as honestly as you can. Your responses will only be visible to the evaluation team at Oxford MeasurEd.

You can find out more about our evaluation and the surveys in the email we sent you inviting you to take part in the survey. You can find out more about how we will use your data in the privacy notice.

If you have any questions or concerns about the evaluation, please contact the Project Director, Lydia Marshall, at mailto:MTPB@oxfordmeasured.co.uk.

Q1 Please enter the unique ID assigned to you. You can find this in the email which contained the link to this survey.

Q2 Are you happy to continue?
Yes >> Q3
No

Thank you – opt out

Thank you for your interest in the evaluation. If you change your mind about taking part in the survey, you can come back using the same link to complete the survey.

CHILD SELECTION

First, we want to ask for your feedback on the child selection task for the MTPB programme.

Q3 Did you receive guidance from the MTPB delivery team on how to use the **Give-N task** to select children for the MTPB programme?

- Yes >> Q4
- No >> Q6

Q4. Did you **follow that guidance** to select the children for the MTPB programme?

- Yes >> Q5
- No >> Why not? (text box on selecting) >> Q6

Q5. How **useful** would you say was the guidance for child selection in selecting children for the MTPB programme?

- Not useful at all
- Somewhat useful
- Useful
- Very useful

Q6. Looking back now, would you say the **right** children in your class were selected for the MTPB programme?

- No, none of the children selected were right for the programme
- Yes, some of the children selected were right for the programme
- Yes, all of the children selected were right for the programme

TRAINING AND SUPPORT

In this section, we want to know about how you found the training and support you received to deliver the MTPB programme.

Q7. What **training/support** did you access? Select all that apply.

- Full-day in-person professional development on 1st March 2024
- Implementation and support for colleagues online session on 7th March 2024
- Online drop-in support session on 21st March 2024
- Half-day in-person professional development on 14th June 2024
- None of these

Q8. How **useful** would you say the training and support from the MTPB delivery team (Fliss James and Melissa Prendergast) has been?

- Not useful at all
- Somewhat useful
- Useful
- Very useful

Q9. How **confident** would you say you felt about **planning the MTPB classroom sessions** after the initial training and support you received from the MTPB delivery team?

- Not at all
- Somewhat
- Quite a bit
- A lot

Q10. How much would you say you **understood** about **how the TA should deliver the MTPB classroom sessions** after the initial training and support you received from the MTPB delivery team?

- Not at all
- Somewhat
- Quite a bit
- A lot

Q11. How **confident** would you say you felt about **carrying out observations and sharing reflections with the TA** after the initial training and support you received from the MTPB delivery team?

- Not at all
- Somewhat
- Quite a bit
- A lot

Q12. Did anything make it **harder** for you to access the **training/support** from the MTPB delivery team? Select all that apply.

- Lack of support from the MTPB delivery team
- Lack of support from your Senior Leadership Team (SLT)
- Lack of resources
- Lack of time
- Other commitments/priorities
- Other, please specify

Q13. Did anything make it **easier** for you to access the **training/support** from the MTPB delivery team? Select all that apply.

- Support from the MTPB delivery team
- Support from your Senior Leadership Team (SLT)
- Resources (e.g., programme handbook, video exemplifications)
- Other, please specify

PLANNING

Next, we want to know about your experiences while planning the MTPB classroom sessions.

Q14. How often did you have **weekly planning sessions** with your TA throughout the 10-week period of the pilot?

- Every week
- At least every second week
- Fewer than every second week
- Only once or twice
- Never

Q15. How much time did you spend on **weekly planning sessions** on average?

- Less than 15 minutes
- 15 – 30 minutes
- 31 – 60 minutes

- More than 1 hour

Q16. Did anything make it **harder** to have the **weekly planning sessions**? Select all that apply.

- Lack of support from Senior Leadership Team (SLT)
- Lack of resources
- Lack of time
- Other commitments/priorities
- Other, please specify

Q17. Did anything make it **easier** to have the **weekly planning sessions**? Select all that apply.

- Support from Senior Leadership Team (SLT)
- Resources (e.g., programme handbook, video exemplifications)
- Other

DELIVERY

Now, we would like to learn about what the classroom delivery of MTPB looked like in your school. We want to know what worked in practice – so please answer honestly.

Q18. How often did **two** MTPB classroom sessions take place in a week throughout the 10-week period of the pilot?

- Every week
- At least every second week
- Fewer than every second week
- Only once or twice
- Never

Q19. Did the MTPB classroom sessions last for **20 minutes** each?

- No, they were less than 20 minutes long
- No, they were more than 20 minutes long
- No, they varied (sometimes less than and sometimes more than 20 minutes long)
- Yes, they were always 20 minutes long

Q20. Did the sessions run in **a dedicated, quiet space** away from the classroom?

- No, never
- Yes, sometimes
- Yes, always

Q21. In the classroom sessions you observed, did the TA deliver the programme using the **Share attention, Respond, Expand, Conversation (ShREC) framework**?

- No, never
- Yes, sometimes
- Yes, always

Q22. Do you think anything made it **harder** for the TA to **deliver** the MTPB classroom sessions? Select all that apply.

- Poor child selection

- Lack of support from Senior Leadership Team (SLT)
- Lack of resources
- Lack of time
- Other commitments/priorities
- Other, please specify

Q23. Do you think anything made it **easier** for the TA to **deliver** the MTPB classroom sessions? Select all that apply.

- Support from you
- Support from your Senior Leadership Team (SLT)
- Resources (e.g., programme handbook, video exemplifications)
- Other

IN-CLASS FEEDBACK AND COACHING

In this section, we would like to learn more about how the teacher-led observation and **in-class feedback and coaching session** took place.

Q24. Did you have **in-class feedback and coaching sessions** with your TA?

- Yes >> Q25
- No >> Q29
- Sometimes >> Q25

Q25. Did the **in-class feedback and coaching session sessions** with the TA take place on the **same day** as the observation you led?

- Yes
- No
- Sometimes

Q26. How much time did you spend on **in-class feedback and coaching sessions** with the TA on average?

- Less than 15 minutes
- 15 – 30 minutes
- 31 – 60 minutes
- More than 1 hour

Q27. Did anything make it **harder** to carry out the **in-class feedback and coaching sessions**? Select all that apply.

- Lack of support from your Senior Leadership Team (SLT)
- Lack of resources
- Lack of time
- Other commitments/priorities
- Other, please specify

Q28. Did anything make it **easier** to carry out the **in-class feedback and coaching sessions**? Select all that apply.

- Support from your Senior Leadership Team (SLT)

- Resources (e.g., programme handbook, video exemplifications)
- Other

OUTCOMES FOR TEACHERS AND TAs

Now, we want to ask you about any potential benefits and consequences for teachers and TAs as a result of taking part in the MTPB programme.

Q29. Has your own **understanding** of how children learn about numbers and operations improved as a result of taking part in the MTPB programme?

- No, not at all
- Yes, somewhat
- Yes, quite a bit
- Yes, a lot

Q30. Has your own **confidence** in teaching maths improved as a result of taking part in the MTPB programme?

- No, not at all
- Yes, somewhat
- Yes, quite a bit
- Yes, a lot

Q31. Would you say your **TA's understanding** of how children learn about numbers and operations improved as a result of taking part in the MTPB programme?

- No, not at all
- Yes, somewhat
- Yes, quite a bit
- Yes, a lot

Q32. Would you say your **TA's confidence** in teaching maths has improved as a result of taking part in the MTPB programme?

- No, not at all
- Yes, somewhat
- Yes, quite a bit
- Yes, a lot

Q33. Were there any **other benefits** for teachers and TAs as a result of taking part in the MTPB programme? Please specify.

Q34. Have you noticed any change in the **relationships** your TA has with the children in your class as a result of their involvement in the MTPB programme?

- No, not at all
- Yes, they have got a bit worse

- Yes, they have got a lot worse
- Yes, they have got a bit better
- Yes, they have got a lot better

Q35. Were there any **negative consequences** for teachers and/or TAs as a result of taking part in the MTPB programme? Please specify.

OUTCOMES FOR CHILDREN

Next, we want to ask you about the potential benefits and consequences for children as a result of taking part in the MTPB programme.

Q36 For how many of the children participating in the programme would you say their **enjoyment** of maths has increased as a result of taking part in the MTPB programme?

- None
- 1 or 2
- 3 or 4
- All of them

Q37. For how many of the participating children would you say their **motivation** towards learning maths has increased as a result of taking part in the MTPB programme?

- None
- 1 or 2
- 3 or 4
- All of them

Q38. Were there any **other benefits** for participating children as a result of taking part in the MTPB programme? Select all that apply.

- Improved social emotional learning
- Improved language and communication
- Improved child-adult relationships
- Other, please specify

Q39. Were there any **negative consequences** for participating children as a result of taking part in the MTPB programme? Select all that apply.

- Time away from other activities
- Stigma of being selected
- Feeling left out for not being selected
- Reduced enjoyment of maths
- Reduced motivation towards learning maths
- Other, please specify

Q40. Do you think any groups of participating children **benefitted more** from the MTPB programme than other participating children?

- Children with English as an Additional Language (EAL) benefitted more than others
- Children with Special Educational Needs (SEN) benefitted more than others
- Children entitled to Free School Meals (FSM) benefitted more than others
- Other group(s) benefitted more than others, please specify
- No group benefitted more

Q41. Do you think any groups of participating children **benefitted less** from the MTPB programme than other participating children?

- Children with English as an Additional Language (EAL) benefitted less than others
- Children with Special Educational Needs (SEN) benefitted less than others
- Children entitled to Free School Meals (FSM) benefitted less than others
- Other, please specify
- No group benefitted less

INTENTION TO SUSTAIN

We are interested in whether you will continue using some of the practices you have learned from the MTPB programme.

Q42. Do you intend to **continue** using any of the following practices now that the MTPB programme has ended? Select all that apply.

- Use of picture books to teach maths
- Use of ShREC technique
- Small-group reading of maths picture books to support children falling behind
- None of these

COSTS

We are now interested in finding out if you, as an individual, incurred any costs as a result of taking part in the programme. Please do **NOT** include the costs incurred by the school, e.g., the cost of a cover while you attended training.

Q43. Did you face any costs for taking part in and delivering the programme?

- Yes >> Q44
- No >> Q46

Q44. Please list the costs you incurred.

Q45. How much did you spend in total?

_____

SCHOOL INFORMATION

Finally, we want to ask some information about how maths is normally taught at your school.

Q46. How **different** is the MTPB programme to how maths is normally taught to reception-age children?

- Very different
- Somewhat different
- Somewhat similar
- Very similar

Q47. Why?

Q48. Is your school currently involved in any **other** early years maths programmes?

- Yes, please specify
- No
- Don't know

Thank you

Thank you for completing this survey. Your answers will be really valuable to the evaluation and we appreciate your time.

If you have any questions, please contact the project director, Lydia Marshall, at mailto:MTPB@oxfordmeasured.co.uk.