



Education  
Endowment  
Foundation

## One Step Ahead

Evaluation report

June 2026

Emma Forsyth, Trisha Harjani, Niall Daly, and  
Patrick Taylor





The Education Endowment Foundation (EEF) is an independent charity dedicated to breaking the link between family income and education achievement. We support schools, colleges, and early years settings to improve teaching and learning for 2–19-year-olds through better use of evidence.

We do this by:

- **Summarising evidence.** Reviewing the best available evidence on teaching and learning and presenting in an accessible way.
- **Finding new evidence.** Funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socio-economically disadvantaged backgrounds. Putting evidence to use.
- **Putting evidence to use.** Supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

We were set-up in 2011 by the Sutton Trust partnership with Impetus with a founding £125m grant from the Department for Education. In 2022, we were reendowed with an additional £137m from government, allowing us to continue our work until at least 2032.

For more information about the EEF or this report please contact:



The Education Endowment Foundation  
5th Floor, Millbank Tower,  
21–24 Millbank,  
London,  
SW1P 4QP



0207 802 1653



[info@eefoundation.org.uk](mailto:info@eefoundation.org.uk)



[www.educationendowmentfoundation.org.uk](http://www.educationendowmentfoundation.org.uk)



Department  
for Education

What  
Works  
Network The logo for the What Works Network, consisting of two large, stylized 'W' characters in a teal color.

## Table of contents

About the evaluator	3	Executive summary	3	Introduction	6	Methods	23	Impact	evaluation																	
42	Implementation and Process Evaluation	53	Conclusion	66	References	72	Appendix A: Security classification of trial findings	75	Appendix B: Effect size estimation	77	Appendix C: Teacher Guide	78	Appendix D: Information and withdrawal form for caregivers	82	Appendix E: Overall guidance for teachers	84	Appendix F: Distribution of outcome measures	85	Appendix G: Regression outputs for primary and secondary analysis	86	Appendix H: Output of missing data analysis	87	Appendix I: Regression outputs and effect sizes for subgroup (FSM) analyses	88	Appendix J: Regression outputs for additional analyses and robustness checks	89

## About the evaluator

This project was evaluated by a team at the Behavioural Insights Team. The project was led by Dr Laure Bokobza and later by Niall Daly, with oversight from Dr Patrick Taylor. Trisha Harjani conducted all quantitative research. The implementation and process evaluation was led by Emma Forsyth with support from Andrés Cueto and Lily Margaroli.

Contact details:

The Behavioural Insights Team  
58 Victoria Embankment,  
London,  
EC4Y 0DS

Email: [info@bi.team](mailto:info@bi.team)

## Acknowledgements

We are grateful to all primary schools that participated in this evaluation, including those who participated in our Teacher Guide pilot. We are particularly thankful to the teachers and teaching assistants who delivered the maths pre-teaching sessions, who completed teacher surveys, and who administered baseline and endline assessments for their pupils. We are also especially grateful to the staff and pupils who as part of our case study visits partook in interviews and focus groups, respectively.

We also thank all staff in schools, academy trusts, and local authorities who promoted the evaluation within their networks, helping us to reach our school recruitment target.

We would like to thank the team at the Education Endowment Foundation, including Jack Mollart-Solity, Ben Sillitoe, Faizaan Sami, Kathryn Davies, Vipin Yadav, and Christine Kelly.

The authors would finally like to give special thanks to colleagues at the Behavioural Insights Team (BIT) who played a key role in project delivery, namely Dr Laure Bokobza, Andrés Cueto, Lily Margaroli, Tess Moseley-Roberts, Hannah Bellier, Rosario Ramos, Laurence Fenn, and Yini Zheng; and also to our expert advisors Dr Kirstin Mulholland (not at BIT) and Lisa Heatherington (not at BIT).

## Executive summary

### The project

The ‘One Step Ahead’ project was designed to evaluate whether a small group, targeted pre-teaching approach could improve mathematical reasoning skills and reduce maths-related anxiety in Key Stage 1 pupils. The trial focused on pupils aged 5 to 7 in Year 1 and Year 2 who were identified by their teachers as likely benefiting from additional support in maths, often because they were working below age-related expectations, experiencing anxiety, or struggling with key concepts. Each participating teacher was provided with a ‘Teacher Guide’ to support their implementation of the practice. The guide gave teachers ideas and practical tools to support them to plan and deliver the pre-teaching sessions with confidence. The guide was developed by the Behavioural Insights Team (BIT) with support from two expert advisors, Dr Kirstin Mulholland and Lisa Heatherington. The intention was for the pre-teaching condition to involve a brief, 10–15-minute session delivered at least twice a week for one term, from September 2024 to December 2024. These sessions, delivered by either the class teacher or a teaching assistant, focused on upcoming lesson content. Pupils in the pre-teaching group received this support in addition to their usual in-class and out-of-class help, while a comparison group received only their usual support with no pre-teaching.

A two-arm randomised controlled trial (RCT) design was used, with randomisation at the pupil level to maximise statistical power. The trial included 2,609 pupils in 222 classes (i.e. 222 pre-teaching groups and 222 comparison groups) across 75 schools. The implementation and process evaluation (IPE) included online surveys for teachers and case studies in six schools, which involved observations, teacher interviews, and pupil focus groups. The trial took place between September 2024 to March 2025. The project was evaluated by BIT.

Table 1: Key conclusions

#### Key conclusions

1. There is no evidence that pupils allocated to receive the One Step Ahead maths pre-teaching made any additional progress in their mathematical reasoning skills, on average, compared to pupils not allocated to receive the pre-teaching. This result has a moderate to high security rating.
2. Among children eligible for free school meals (FSM), the best estimate of pre-teaching’s impact is a small negative effect on pupils’ mathematical reasoning skills. These results may have lower security than the overall findings because they relate only to a subsample of pupils.
3. Teachers faced challenges with timetabling and staff availability, with many reporting difficulty fitting sessions into their daily schedule. This led to inconsistent delivery and sessions that were often shorter than the recommended 10–15 minutes. School size may have been a factor here, with some teachers in multi-form entry schools noting particular challenges in timetabling.
4. Teachers perceived pre-teaching to have a positive effect on pupils’ confidence and engagement, teachers attributed this to small groups’ low stakes environment, teachers building explicit links between the pre-teaching and the main lesson (‘bridge building’), and changes in how the wider class perceived the pupils in the group. These perceptions from teachers were not reflected in the impact evaluation outcomes of maths attainment and maths related anxiety.
5. The pupil selection process also posed challenges, and randomisation sometimes led to pupils who teachers felt would not benefit most from pre-teaching (i.e. they were working at or above expectations) being included in the pre-teaching group. This meant that some pupils were included in the pre-teaching group even though they were not the intended target, resulting in a mixed-ability group. These factors made it harder to tailor sessions effectively.

### EEF security rating

These findings have a moderate to high security rating. This was a Teacher Choices trial, testing an everyday choice teachers make. The trial was well-powered. However, the following factors reduced the security of the trial. 16% of the pupils who started the trial were not included in the final analysis. There were some important differences in prior attainment between the pupils in treatment and comparison groups. There were challenges to delivering pre-teaching with fidelity and outcome measures were teacher administered, increasing the possibility of bias. There is also some evidence that pupils in the control group received pre-teaching which may have reduced the overall effects observed.

## Additional findings

There is no evidence that pupils allocated to receive the One Step Ahead maths pre-teaching made any additional progress in their mathematical reasoning skills, on average, compared to pupils not allocated to receive the pre-teaching. This is our best estimate of impact, which has a moderate to high security rating. As with any trial, there is always some uncertainty around the result: the possible impact also includes a small negative effect and a small positive effect.

While teachers reported perceived improvements in pupil confidence and engagement through ‘bridge building’ and a low stakes environment, the impact evaluation did not find evidence of positive impact on pupil outcomes, either for mathematical reasoning or maths-related anxiety. This may reflect implementation challenges rather than the effectiveness of the approach itself. Inconsistent delivery, driven by timetabling constraints and staff capacity, appears to have reduced the dosage and quality of sessions. In addition, pupil groups were not always well targeted, which made it harder to tailor support, and some teachers adapted sessions away from the intended ‘forward-looking’ focus of the sessions by incorporating ‘catch-up’ elements.


While using teaching assistants helped some schools to ensure sessions were delivered, teaching assistants delivering them might have moderated the impact of pre-teaching in some instances, and could help to explain the lack of treatment effect. In interviews, some school project leads felt that teaching assistant-led pre-teaching could be less effective than teacher-led sessions. The Education Endowment Foundation (EEF) guidance does not rule out using teaching assistants to deliver well-chosen, evidence-based, structured programmes where appropriate, but it may be that this particular type of approach is not appropriate for teaching assistants, especially in resource-constrained schools, where they may not receive the support needed to deliver it well (EEF, 2025a).

The trial’s findings have implications for whether pre-teaching can be effectively replicated in normal practice conditions. While the choice was designed to be low cost and low burden, the IPE revealed that even these requirements were difficult for some schools to sustain. Some teachers cited a lack of staff time and resource constraints as the biggest barriers to consistent delivery. Reports of shorter than intended sessions and difficulty in consistent timetabling suggest that the intended dosage was not always met. Interviews with teachers highlighted that increasing the length of the sessions to provide a higher dosage would have been difficult to fit into their daily schedules, and while an alternative way of increasing dosage might be delivering sessions less frequently over a longer period, some teachers felt pre-teaching needed to be frequent and consistent.

In summary, the findings contribute to the evidence base by highlighting the importance of implementation fidelity for non-prescriptive teacher choices. The IPE findings suggest that a lack of protected time was a key barrier to consistent delivery, offering a plausible explanation for the lack of detected impact on mathematical reasoning. While some teachers perceived a boost in pupil confidence, the impact evaluation found no reduction in maths anxiety. This suggests that even low-cost strategies require protected time and careful planning to be effective. It also builds on existing evidence that pre-teaching can be beneficial for building confidence, enhancing content retention, and reducing anxiety, but is contingent on effective implementation, frequency, and ‘intensity’ of the model (van den Berg *et al.*, 2017).

## Impact

Table 2: Summary of impact on primary outcome(s)

Outcome / group	Effect size (95% confidence interval) Hedges' g	The EEF security rating	No. of pupils	P-value
Mathematical reasoning skills (all pupils)	0.01 (-0.07, 0.09)		2,186	0.80
Mathematical reasoning skills (FSM pupils)	-0.09 (-0.25, 0.07)	N/A	610	0.28

## Introduction

### Overview of evaluation design

The Education Endowment Foundation (EEF) commissioned the Behavioural Insights Team (BIT) to evaluate a 'Teacher Choice' focused on maths pre-teaching in Key Stage 1, called One Step Ahead. This evaluation was motivated by the policy context of raising academic standards and improving teaching approaches in core subjects like maths, which is a key priority for the Department for Education (DfE) (DfE, 2025a). Mathematics also plays a crucial role in pupils' development, and a pupil's early mathematical understanding has been shown to correlate strongly with later educational achievement (Duncan *et al.*, 2007; Jordan *et al.*, 2009). The disadvantage gap for overall attainment has increased in the past five years, from 9.3 to 10.3 months at the end of primary school between 2019 and 2022 (EPI, 2023). Evidence shows that additional targeted support in maths can be highly beneficial, but there is a gap in evidence when it comes to the best way to provide out-of-class support in maths, specifically for young pupils (EEF, 2021). One Step Ahead aims to fill this gap and generate evidence that can be easily and effectively used by teachers to work towards closing the attainment gap in Key Stage 1 maths.

The theoretical rationale for pre-teaching is that providing targeted support to small groups of pupils before a lesson can improve their confidence, reduce anxiety, and help them master key concepts and vocabulary (Lalley and Miller, 2006; Watt and Therrien, 2016). This, in turn, is hypothesised to improve their mathematical reasoning skills (EEF, 2020) and their engagement in the subsequent main maths lesson(s) (Trundle, 2018). This core rationale relies on forward-looking practice, specifically previewing upcoming lesson content and introducing new vocabulary and concepts to enable access to the main lesson. This is distinct from remediation or catch-up, which focuses on consolidating or revisiting past content that the pupil may have previously failed to grasp (EEF, 2021). While some research suggests that the frequency and 'intensity' of pre-teaching correlate with attainment outcomes (van den Berg *et al.*, 2017), the overall low prevalence of the practice in UK schools (see for instance the implementation and process evaluation [IPE] findings, with 24% of surveyed teachers reporting they do so regularly or occasionally) meant there was a need for a trial to generate evidence on its impact.

The initial scoping phase, conducted from October 2022 to January 2023, laid the groundwork for the evaluation by investigating four potential research questions for a Teacher Choices trial. This phase included a rapid evidence review, 12 semi-structured interviews with educators, and a large-scale online survey of 743 Key Stage 1 teachers via the Teacher Tapp app. The purpose of this research was to gauge current practices, identify genuine dilemmas for teachers, and assess the feasibility and potential impact of a trial. Initially, four research questions were considered, focusing on topics such as pre-teaching versus post-teaching, timing of interventions, use of manipulatives, and teacher modelling. The decision to focus on the pre-teaching question was influenced by feedback from the EEF and the Study Advisory Board, who prioritised questions with high teacher interest and potential for social impact.

Further in-depth scoping on pre-teaching confirmed its suitability for a trial. While the practice of regular, forward-looking, small group pre-teaching was not widespread, with only 24% of surveyed teachers reporting that they did so regularly or occasionally, teachers expressed a high level of interest in the topic and the potential evidence a trial would generate. This interest was particularly strong among those who wanted to know how to implement pre-teaching effectively and assess its impact on pupil attainment and confidence. The scoping also revealed that teachers viewed the decision of whether or not to pre-teach at all as a significant choice they face in their practice. This led to the design of a trial comparing pre-teaching to a business as usual approach, aiming to provide clear evidence to support teacher decision-making.

To address these questions, the study employed an integrated evaluation design featuring a two-arm randomised controlled trial (RCT) and a mixed-methods implementation and process evaluation (IPE). Randomisation at the pupil level was chosen to maximise statistical power. The RCT compared the impact of providing teachers with guidance on delivering regular pre-teaching sessions in maths with a business as usual approach. The IPE, which included surveys, case studies, and qualitative interviews, was designed to explain the impact evaluation findings by exploring the causal links outlined in pre-teaching's logic model, and identifying barriers and facilitators to implementation. This combined approach provides both a measure of the choice's impact and an improved understanding of how and why the outcomes occurred.

## Teacher Choice approaches

### Description of the choice

Table 3 below gives an overview of the choice that was evaluated. The choice was developed based on the findings from the scoping research and the support of two experts who supported BIT to develop the guidance and conduct a pilot evaluation after the scoping phase and before the main trial.

- **Dr Kirstin Mulholland.** Former primary school teacher and leader, and a former maths content specialist for the EEF. She was a member of the Study Advisory Board for this project in its initial stages. She currently works as an assistant professor in Education at Northumbria University, with a specialism in innovative approaches for professional development, teacher research, perceptions and experiences of teaching and learning, and the role of student voice in pedagogy development.
- **Lisa Heatherington.** Former primary school teacher and leader, the National Centre for Excellence in the Teaching of Mathematics (NCETM) accredited professional development lead, and assistant lead at the Great North Maths Hub. Her specialism lies in teacher Continuing Professional Development (CPD), with a focus on creating meaningful learning and teaching. She led working groups with teachers to implement systematic pre-teaching as part of her work with the Great North Maths Hub.

Table 3: Choice description

	Choice A Pre-teaching condition	Choice B No pre-teaching business as usual condition
Structural and logistical factors	Regular (two or more times a week), short pre-teaching sessions delivered to a targeted small group of pupils before the main lesson, for one term	No out-of-class targeted support for pupils before the main lesson during one term
Other forms of support to pupils	Unchanged	Unchanged
Hypothesised mediators	Active connections and bridge-building; improved retention of key concepts and vocabulary; increased engagement and confidence; reinforcement of lesson-critical vocabulary and concepts	Unspecified—depends on what business as usual looks like in each class

Existing evidence and findings from the scoping phase allowed BIT to identify some core elements for the pre-teaching condition as we define it. To ensure that teachers in the ‘Choice A’ arm delivered the kind of pre-teaching that we anticipated would have an impact and sought to evaluate, we provided them with a guidance document (see [Appendix E](#)), referred to as the ‘Teacher Guide’ in this report. In the guide, teachers are asked to plan and deliver the pre-teaching sessions following these requirements:

1. **Frequency.** At least twice a week.
2. **Timing.** Before the main lessons (focusing on upcoming lesson or lesson unit content).
3. **Length.** 10–15 minutes.
4. **Duration.** One term.
5. **Recipients.** Fixed, targeted small group of six pupils, identified based on teacher’s anticipation of pupils’ needs.
6. **Person delivering.** Class teacher if possible and/or teaching assistant otherwise.
7. **Key aim.** Enable access to the main lesson.<sup>1</sup>

This definition of pre-teaching captures the core components of the practice in the logic model, validated by the experts who supported BIT during the scoping phase, Dr Kirstin Mulholland and Lisa Heatherington (see Figure 1 for the full logic model). Adopting this pre-teaching practice was deemed acceptable and feasible by the teachers and experts that we

<sup>1</sup> This is essential as it distinguishes pre-teaching from other forms of targeted support, including remediation and other forms of responsive support programmes.

consulted throughout the scoping phase and Teacher Guide pilot. Specifically, we assessed teachers' views about the acceptability and feasibility of implementation of the practice during a co-design workshop during the scoping phase, and during the feedback and debrief workshop at the end of the pilot. Dr Kirstin Mulholland, Lisa Heatherington, and the Study Advisory Board member Claire Williams also validated this stance. The choice of a duration of one term was motivated by the trade-off between having a long enough period to observe a change in pupils' outcomes, while enabling the group of pupils initially identified as needing additional support to remain fixed, since pre-teaching is designed to be adaptive and responsive to pupils' needs at a specific point in time. An additional consideration was the time over which teachers would be able and willing to comply with the requirement to deliver the sessions twice a week or more, to a fixed group of pupils.

Each core component has a clear impact-maximising justification, building on existing evidence showing that frequent programme over sustained periods of time are more effective than ad hoc support (Kroesbergen and Van Luit, 2003; Slavin *et al.*, 2017), that having a small, fixed group of pupils is important to monitor pupil progress and tailor instruction effectively (Vaughn *et al.*, 2003, Wanzek and Vaughn, 2007), and that teacher-led programmes tend to yield better outcomes due to their familiarity with the pupils allowing them to tailor the support (Slavin *et al.*, 2017, Wanzek and Vaughn, 2007). In addition, we designed these core components so as to give teachers sufficient flexibility in what they pre-teach and in how they go about it, which teachers and maths leads said was important in the scoping phase. This aimed to help ensure that teachers could meet the specific needs of their pupils and operate within their practical constraints while delivering the pre-teaching sessions as intended.

We identified other, non-core practice components in the scoping phase and during the development of the logic model as likely to boost the impact of the pre-teaching. We did not require teachers to commit to implementing these. Instead, we highlighted them in the guidance and allowed teachers to adapt their practice based on their constraints and pupils' needs. The impact-boosting elements are:

1. **Scheduling.** As close as possible to the main lesson (for instance, the morning of or the day before).
2. **Strategies and activities for pre-teaching.** Given as examples in the guidance.
3. **Strategies to make 'active links' between pre-teaching and main lesson.** Given as examples in the guidance.

Table 4 below summarises the above and provides a detailed description of the choice.

Table 4: Choice description

Name	Pre-teaching	Comparison (no pre-teaching business as usual)
Why (Rationale)	Targeted additional support focusing on upcoming lesson or lesson unit content has the potential to improve pupils' mathematical reasoning skills and confidence	No targeted additional support before the main maths lesson, but all other pre-existing support during and outside of lesson time should take place as usual
Who (Recipients)	Six pupils in each Year 1 and Year 2 class in participating schools, who have been identified by teachers as having the potential to benefit from additional targeted support in maths prior to the main lesson, and were randomly assigned to the pre-teaching group	Six pupils in each Year 1 and Year 2 class in participating schools, who have been identified by teachers as having the potential to benefit from additional targeted support in maths prior to the main lesson, and were randomly assigned to the comparison group
What (Materials)	A guidance document explaining the basic requirements that teachers must follow, and including examples of strategies and activities that they could use  No restriction on the content or structure of the pre-teaching sessions. Teachers are left free to adapt the content and structure depending on their lesson plans and pupils' specific needs	None other than the instruction not to pre-teach during the trial period  All other pre-existing support during and outside of lesson time should take place as usual

	All other pre-existing support during and outside of lesson time should take place as usual	
What (Procedures)	Pre-teaching sessions in maths (in a pre-teaching group of six pupils), focusing on upcoming lesson or lesson unit content	No restriction other than no pre-teaching
Who (Practitioners)	The sessions should ideally be delivered by the teacher, but can be delivered by a teaching assistant if needed	Not applicable
How (Delivery guide)	Teachers received a Teacher Guide to support them to plan and deliver the pre-teaching sessions (see <a href="#">Appendix C</a> for details)	
Where (Location)	In schools	Not applicable
When and how much (Dosage)	10–15-minute pre-teaching sessions, at least twice a week, for one term (September to December 2024)	No restriction other than no pre-teaching, for one term (September to December 2024)
Tailoring (Adaptation)	Teachers are encouraged to adapt the content and structure of the pre-teaching sessions to: i) their lesson plans; and ii) pupils' specific needs	Not applicable

## Logic model and narrative

Figure 1 below presents the logic model for the choice that was evaluated. The logic model was developed iteratively. We produced a first draft during an internal workshop involving only the BIT researchers on the project during the scoping phase. We then developed the accompanying narrative and shared it along with the draft logic model with Dr Kirstin Mulholland and Lisa Heatherington. We gathered their feedback and input during a one-hour online session, following which we refined the model. We also sought and incorporated the EEF's feedback on the model. The rest of this section describes the key components of the model in more detail.

### Inputs

#### *Teacher Guide*

Each participating teacher was provided with a 'Teacher Guide' to support their implementation of the practice (see [Appendix C](#) for the full Guide). The guide introduced the study's objectives and gave participating teachers ideas and practical tools to support them to plan and deliver the pre-teaching sessions with confidence for one term. The guide included:

1. A planning and delivery checklist.
2. Key tips and guidance.
3. A list of do's and don'ts for planning and delivery.
4. Illustrated examples of session structure and activities.
5. A worked example.
6. Frequently asked questions (FAQs).
7. A list of freely available resources.

We expected engagement with the guidance to vary by individual teachers depending on a range of factors, including general teaching experience, experience with pre-teaching, knowledge of pre-teaching strategies, and experience of delivering small group targeted support. Some teachers may have relied more on the guidance if they did not feel as confident delivering pre-teaching independently. We also encourage teachers to exercise their professional judgement in the structure and delivery of the sessions. A first version of the guide was tested with six teachers in January 2024. We collected their feedback throughout and at the end of the implementation period via surveys and a debrief workshop. This led to iterations and improvements for the version of the guide used during the trial.

#### *School-level planning*

Participating schools and teachers were asked to conduct some planning activities as part of the project. This senior-level logistical support was identified as a key enabler of successful implementation in the scoping phase and Teacher Guide pilot. This included timetabling the sessions at least twice a week, ideally delivered by the class teacher but delivered by another teacher or a teaching assistant if necessary, ensuring cover for the rest of the class during this time. Schools were also asked to find a space suitable for delivering the sessions. We expected Year 1 and Year 2 teachers to collaborate with

their senior leadership team (SLT) to ensure delivery was feasible. We anticipated that this may have varying effects on delivery of the pre-teaching in individual schools, based on different interrelated factors such as school size, class sizes, senior buy-in for the Teacher Choice, and availability of resources. To support this, we required that an SLT member sign the Memorandum of Understanding (MoU) when a school signed up for the project. We also assigned a 'project lead' for each school, who was often also an SLT member and could also be a teacher delivering the pre-teaching themselves, and was in charge of coordinating implementation and supporting teachers throughout the trial.

As is detailed in the 'Methods' section, each participating class teacher identified 12 pupils in their class as eligible to receive the pre-teaching, and if these children's parents/caregivers did not withdraw them from the study, these pupils were randomised into a pre-teaching group (of six pupils) and a comparison group (of six pupils). We provided no strict guidance on pupil 'eligibility' for teachers, other than suggesting that the 12 pupils put forward in each participating class were 'likely to benefit from pre-teaching in maths', and leaving it up to teachers to apply this to their own pupils.

## Activities

The main activity of the One Step Ahead Teacher Choice involved delivering two or more pre-teaching sessions per week in Key Stage 1 maths for one term (September to December 2024). The core components of the practice are defined as follows:

1. Around 10–15-minute pre-teaching sessions.
2. Delivered before the main lesson.
3. To a fixed group of pupils (six pupils).
4. Two or more times a week.
5. Delivered by a teacher, teaching assistant, or a combination of both.
6. Focusing on upcoming lesson content.

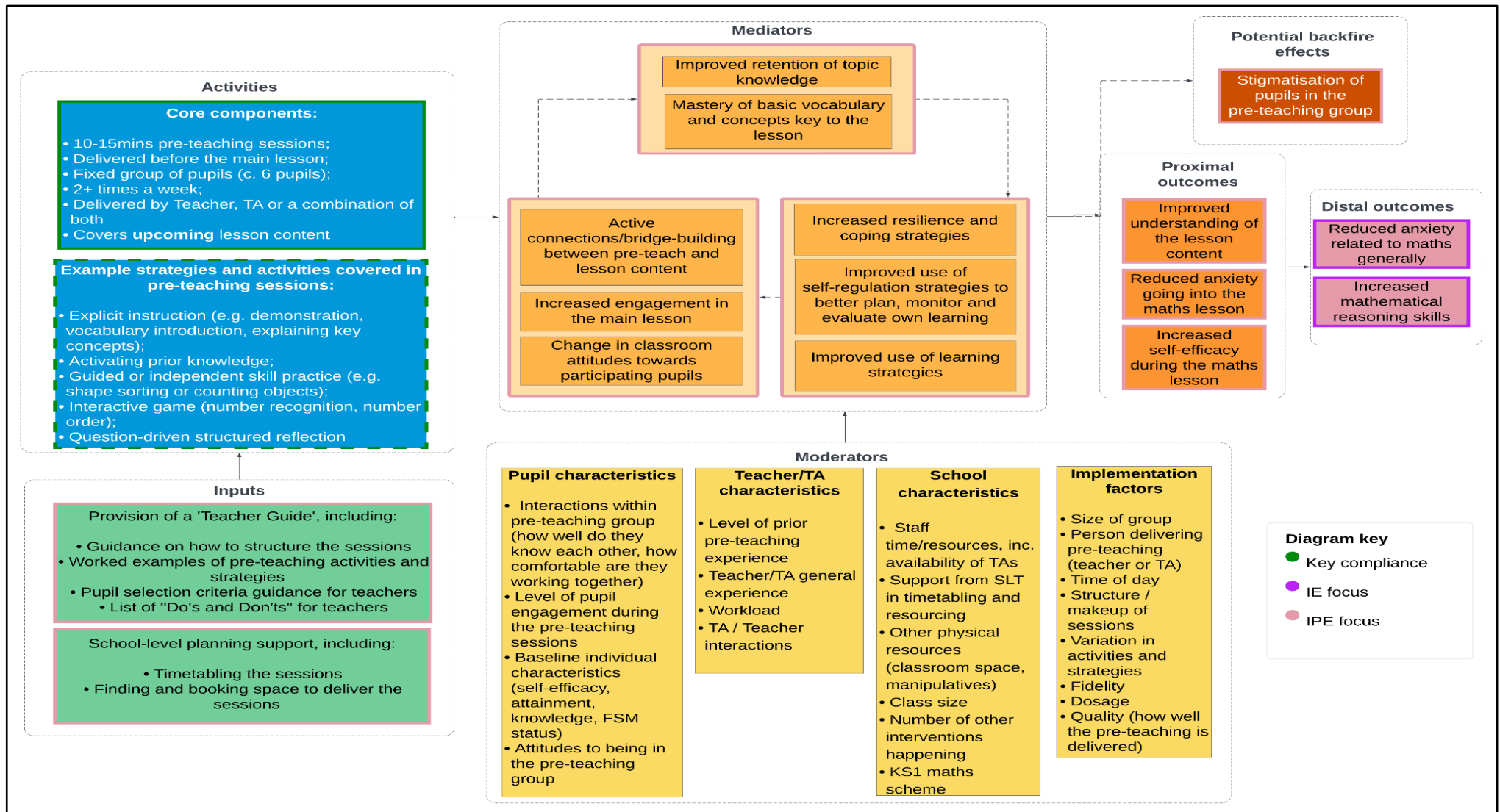
These core components identified in the logic model were fixed. Teachers were asked to comply with them in order to ensure a minimum level of fidelity to the model of pre-teaching, which we believed could have the biggest positive impact on pupils. We refer to those as 'key compliance' factors in the logic model.

By contrast, the structure and content of the pre-teaching sessions was not prescriptive, teachers were largely free to structure, plan, and deliver the sessions as they wanted. This flexibility was deemed necessary to account for the fact that teachers would be teaching different topics in a different order during the implementation period. Scoping stage findings showed that being prescriptive about this would have reduced acceptability and compliance with the practice. The guidance provided to teachers suggested that the structure and content of the lessons should be driven by upcoming lesson content (although not necessarily limited to the immediate upcoming lesson), and pupils' specific needs, which may evolve over time. This is in line with findings from the scoping, existing research on targeted support, and the recommendations of our expert advisors to maximise the likelihood of impact.

With input from our expert advisors during the guidance development and feedback from teachers during the pilot, we provided a non-exhaustive and illustrated list of key example strategies and activities that teachers can use in the Teacher Guide, including:

- interactive games;
- the use of virtual representations and manipulatives;
- addressing misconceptions;
- storytelling;
- prediction and discussion;
- scaffolding;
- planning and teaching;
- making comparisons; and
- boosting mathematical literacy.

Figure 1: Pre-teaching logic model



The guidance encouraged teachers to pick and choose activities and strategies from the list of examples. We acknowledge that the extent to which teachers engaged with these likely depends on various factors, including their experience and/or knowledge of pre-teaching and small group instruction more broadly. We aimed to investigate this in the IPE where we asked teachers about their use of the Teacher Guide and views on this.

## Mediators

We define mediators as the mechanisms or causal links through which the pre-teaching activity is expected to produce the desired outcomes. By contrast, moderators are the factors (such as pupil, teacher, or school characteristics) that influence the strength or direction of the relationship between pre-teaching and the outcomes. As there was some uncertainty at the time of drafting the Teacher Guide over how reliably teachers would be able to deliver or mobilise each of the mediators outlined in the logic model, we adopted a maximalist approach (including all key plausible mediators in the logic model), while recognising that some may be more influential than others. We grouped the mediators into sub-categories, depending on the proximal outcomes to which they are most directly linked (maths anxiety and self-efficacy during the lesson, and understanding of lesson content). However, we believe many of these mediators are contingent upon each other, and as a result may feed into more than one outcome. This interdependence of mediators means that the RCT is not designed to statistically isolate the effect of any single mechanism. Instead, the IPE was specifically designed to explore and triangulate these contingent causal links qualitatively, observing which mechanisms were most activated and how they interplayed to influence perceived outcomes.

### *Mediators related to maths anxiety and self-efficacy*

1. Improved use of self-regulation strategies to better plan, monitor, and evaluate own learning.
2. Increased resilience and coping strategies.
3. Improved use of learning strategies.

Pre-teaching has the potential to reduce pupils' feelings of anxiety related to maths and improve their self-efficacy by improving their capacity to effectively deal with challenges, stress, and adversity in the learning environment. Pre-teaching can also contribute to equipping pupils with specific methods to manage stress and adversity (e.g. through seeking help from teachers or peers and developing problem-solving skills and improved emotional regulation). In addition, pre-teaching can help pupils gain a better understanding of and ability to use a range of learning strategies, namely, techniques to acquire, process, and retain information more effectively.

These mediators may be contingent on teachers employing self-regulation strategies as an activity in the sessions. However, should teachers implement this kind of strategy in delivering pre-teaching sessions, evidence suggests they develop confidence in their ability to engage with mathematical concepts (De Corte *et al.*, 2011). Should they get 'stuck' or struggle with content or concepts, they learn strategies and coping mechanisms to help them deal with this in the main lesson itself. Furthermore, the success of the mediators may also be influenced by the pedagogy of the main lesson itself. If the main lesson actively reinforces or provides opportunities to apply pre-taught strategies, the impact on confidence and coping mechanisms is likely magnified. Relatedly, it is worth noting that if a school's or teacher's approach to maths lessons is already highly differentiated (i.e. different groups doing different tasks based on ability), the introduction of pre-teaching sessions may not have as much impact on recipient pupils as it might do otherwise.

In addition, we identified some mediators related to how pre-teaching can affect what happens during the main maths lesson:

- active connections/bridge-building between pre-teach and lesson content;
- increased engagement in the main lesson; and
- classroom attitudes towards participating pupils change.

Our expert advisors suggested—based on their experience with pre-teaching and their expertise—that teachers building ‘active links’ between the pre-teaching and the main lesson helps reduce maths-related anxiety among pupils. This refers to the practice of teachers purposefully building links between the pre-teach and the subsequent lesson(s). This can be done in a number of ways, for example, by reminding pupils about the content and activities already covered in the pre-teaching session, or by highlighting achievements or progress that pupils made during the pre-teach, in the main lesson. The idea is to bridge the gap between the pre-teaching and the main lesson.

This reinforces the message to the pupils that they ‘can do’ maths. It may also shift the attitudes of other pupils in the class as to who is ‘good’ or ‘bad’ at maths, thereby reinforcing the confidence of pupils in the pre-teaching group. Participation in the pre-teaching sessions can also give them more confidence to participate in class discussions in the main lesson, as they have had a ‘preview’ of concepts or vocabulary that may come up ahead of time, and therefore, will feel less worried about giving the wrong answers. We anticipated that this confidence building would have the effect of reducing anxiety related to maths, especially when being exposed to the pre-teaching repeatedly and regularly over the course of one term.

#### *Mediators related to understanding lesson content*

1. Improved retention of topic knowledge.
2. Mastery of basic vocabulary and concepts key to the lesson.

The assumption underpinning these two mediators was that by introducing key concepts and vocabulary ahead of the lesson, pupils will have more chances to understand them and would not get stuck as easily in the main lesson. By getting access to this content in small groups ahead of time, pupils would have more time to discuss uncertainties and unpick misconceptions with the teacher before they arise in the main lesson. Overall, we hypothesised that regular pre-teaching would enable an improved understanding and retention of lesson content over time, through a better grasp of mathematical language and concepts, enabling better problem-solving and more effective communication.

#### **Outcomes**

We identified three proximal outcomes for pupils. They are the observable positive changes that we expected pupils to experience in the main lesson. Due to either difficulty in finding or using a suitable measure for Key Stage 1 pupils, and/or due to the need to minimise the assessment burden on both pupils and school staff, we did not directly measure these outcomes in the impact evaluation.

1. **Improved self-efficacy in maths during the main lesson:** Self-efficacy is an individual’s belief in their ability to perform a task, and to achieve the desired outcome by doing so (Bandura, 1997).
  - a. *Rationale.* Having covered some of the upcoming content or practiced some skills in the pre-teaching session, pupils would be more likely to believe in their mathematical ability to do what is required of them and to get the correct answers. Early exposure to concepts and tailored support helps build pupils’ confidence.
2. **Reduced maths anxiety during the main lesson:**
  - a. *Rationale.* As they would have had a preview of the main lesson’s content, pupils would experience less anxiety linked with uncertainty over the upcoming maths lesson. This increased familiarity and comfort can contribute to a more positive learning experience and decrease feelings of stress.
3. **Improved understanding of lesson content:**
  - a. *Rationale.* Pre-teaching can help pupils unlock a better understanding of the mathematical concepts and language covered in the upcoming main lesson. Repeated exposure to pre-teaching should enable pupils to have a better grasp of mathematical concepts and language. A stronger grasp of these concepts enables pupils to develop their problem-solving skills and make connections between different concepts and areas of maths. We hypothesised that improved mastery of mathematical language

through pre-teaching would help pupils to understand problems and instructions, and to articulate their thinking.

These proximal outcomes are directly related to the distal outcomes, which we expected to accrue over the longer run. These are the outcomes, which we sought to measure in the impact evaluation.

#### 1. Improved mathematical reasoning skills:

- a. Rationale. In the long run, sustained improved understanding of lesson content contributes to improving pupils' reasoning skills, via the mediators outlined above. Mathematical reasoning skills are defined as the foundational abilities that lay the foundation for pupils' mathematical learning. It includes the ability to compare quantities and sizes, use logical thinking to explain ideas, understand spatial relationships, recognise patterns, and grasp fundamental numerical concepts. These skills start to develop before pupils start to learn arithmetic in school (Nunes *et al.*, 2015).

#### 2. General reduced anxiety related to maths:

- a. Rationale 1. We hypothesised that by improving pupils' confidence in their own mathematical abilities, prolonged additional support in the form of pre-teaching would help reduce the general anxiety that pupils experience in relation to maths learning.
- b. Rationale 2. By helping them believe that they 'can do' maths, pupils would feel more comfortable engaging and participating in maths lessons more generally, over the course of their education. This increased sense of competence in turn creates a more positive learning experience.

Some evidence suggests a positive relationship between improved self-efficacy and reduced anxiety and mathematical abilities (Hembree, 1990). However, the direction and strength of a causal relationship has been subject to debate, and recent research highlights the complex interactions between psychological factors and mathematical achievement (Krinzinger *et al.*, 2009). For this reason, we did not specify a causal relationship between the two distal outcomes.

A potential negative unintended consequence of the trial was that pupils who were selected for the pre-teaching are to some extent stigmatised by their peers. This could have had a detrimental effect on their well-being, as well as their confidence and attainment in maths. We note this as a potential backfire effect in the logic model. We discussed this issue with Dr Kirstin Mulholland and Lisa Heatherington, as well as some Study Advisory Board members and teachers engaged during the scoping phase, which led us to consider this risk as low. Key Stage 1 pupils are regularly taken out of class and split into groups during the course of the day/week to be provided with additional support, or due to resources and scheduling in the school. It is therefore, unlikely that for pupils this age, getting some additional support within regular school hours would be a source of stigma.

To mitigate this risk, the Teacher Guide encouraged teachers to come up with a fun name for the pre-teaching group (e.g. 'Sneaky Peaks') and to build connections between the pre-teaching session and the main lesson in a way that highlights pupils' achievements.

### Issues arising

It is important at this stage to note a number of issues arose during the implementation of the evaluation, which may have impacted the trial's internal validity and the generalisability of the findings. These issues relate primarily to implementation, and data collection.

#### *Baseline assessment and choice implementation*

A key challenge emerged early in the implementation phase related to teacher workload and school capacity. The baseline assessment period was initially set for the first three weeks of the Autumn Term but proved challenging for some schools to complete within this time frame due to high workloads and the time required for one to one test administration for younger pupils. Further detail on this challenge is presented in the '[Trial design](#)' and '[Subgroup analyses](#)' sections in the 'Methods' section below.

Additionally, several other implementation factors may have affected the choice's impact. Teacher surveys revealed significant challenges with timetabling and staff availability, cited by 72% and 56% of respondents, respectively. This sometimes resulted in sessions being shorter than the recommended 10–15 minutes and may have affected delivery consistency. The pupil selection process also proved difficult, with selection criteria that sometimes led to mixed-ability groups that may not have been as targeted as intended. Some teachers felt the trial's requirement to nominate 12 pupils per class was too high, as they did not have that many pupils who genuinely needed the support. Consequently, after random allocation, the six-pupil pre-teaching groups sometimes included pupils who teachers would not have typically selected for this type of targeted approach, in the absence of a trial. Further detail is provided in the IPE findings, which argue that this was a moderating factor, making sessions difficult to tailor and potentially diluting their impact. Qualitative feedback also indicated some teachers deviated from the intended 'forward-looking' focus of the sessions by incorporating 'catch-up' elements, which was a clear deviation from the guidance provided and the idea of pre-teaching.

### **Data collection and measurement**

To maintain consistency in measurement of our primary outcome (mathematical reasoning skills), the Quantitative Reasoning Test (QRT), was administered by teachers for both year groups. We worked with the assessment's developer, Dr Terezinha Nunes, and the Reasoning First team to tailor the assessment to the trial's aims. Dr Terezinha Nunes recommended administering the test in a one to one format for Year 1 pupils, and in a whole-group format for Year 2 pupils. The one to one tests were designed to take approximately 30 minutes to complete per pupil. In contrast, for Year 2 assessments, 45 to 60 minutes was expected to be sufficient time for the assessment delivery to the whole group (12 pupils per class). However, teacher-administered outcome tests presented another challenge. Around 65% of endline survey respondents found test administration difficult, citing it as time-consuming, particularly the one to one assessments for Year 1 pupils. This contributed to delays in data submission. Additionally, teachers raised concerns that the primary outcome measure, the QRT, was too difficult and not sufficiently aligned with their curriculum. A total of 423 pupils (16.2%) were excluded from the primary analysis due to withdrawals or missing test results, a slightly higher rate than the originally assumed 14%.

## Evaluation objectives

### Impact evaluation (IE) design

#### Research questions

The overarching research question we seek to address is: Is small group, targeted pre-teaching effective in improving pupil mathematical reasoning skills, above and beyond business as usual targeted support in Key Stage 1? This research question reflects the fact that we are not restricting the other forms of support that pupils who are assigned to the pre-teaching group receive, both outside of and during the main lesson. In other words, we are not asking teachers to change their practices other than to incorporate the pre-teaching for the treatment group.

As per the evaluation's published [study plan](#) (Bellier *et al.*, 2024), the **primary research question** of the impact evaluation is the following:

1. Was there a difference in **pupils' mathematical reasoning skills** at the end of the Autumn Term (as measured by the QRT) between pupils who were allocated to receive the pre-teaching, and pupils in the comparison group?

Additionally, the impact evaluation aimed to address the following **secondary research questions**:

2. Was there a significant difference in **pupils' anxiety related to maths** at the end of the Autumn Term (as measured by their score on the Children's Mathematics Anxiety Scale UK [CMAS-UK]), between pupils who were allocated to receive the pre-teaching and pupils in the comparison group?
3. Are there heterogeneous effects by free school meals (FSM) status? By baseline attainment (for pupils' mathematical reasoning skills only)?

As pupils' anxiety related to maths was not measured at baseline, we only tested for heterogeneous effects using baseline mathematical reasoning skills.

#### Implementation and Process Evaluation (IPE)

The IPE addressed the following research questions:

1. How are pre-teaching sessions implemented by teachers, and what variations exist? (*fidelity*)<sup>2</sup>
  - 1.1. What did the pre-teaching sessions look like, when were they delivered, and for how long? To what extent did teachers stick to the core components? (*adherence*)
  - 1.2. How much pre-teaching did pupils in the pre-teaching group receive? (*dosage*)
  - 1.3. How well was pre-teaching delivered? (*quality*)
  - 1.4. What adaptations were made and why? (*adaptations*)
2. How does pre-teaching differ from schools' usual practice of providing targeted support for Key Stage 1 maths? (*choice differentiation*)
  - 2.1. What business as usual support did pupils in the pre-teaching and comparison group receive?
  - 2.2. To what extent did schools comply with the core restriction to not provide support before the lesson to pupils in the comparison group? In cases where schools did not comply, what were the drivers of non-compliance?
3. What factors influence the implementation of pre-teaching, and in what ways? (*moderators*)
  - 3.1. What characteristics of the school environment influence implementation, and in what ways?
  - 3.2. What teacher characteristics influence implementation, and in what ways?
  - 3.3. What pupil characteristics influence implementation, and in what ways?

---

<sup>2</sup> Fidelity has five core components (Carroll *et al.*, 2007): adherence, exposure, or dosage; quality of delivery; participant responsiveness; and programme differentiation.

- 3.4 What characteristics of the pre-teaching guidance influence implementation, and in what ways? Which implementation requirements (if any) would teachers change if they adopted pre-teaching outside of the trial?
4. What are the perceived outcomes of pre-teaching for pupils, including unintended outcomes? (*perceived impact*)
  - 4.1 What are the perceived outcomes of pre-teaching for pupils according to teachers?
  - 4.2 How did pupils feel the support they received prior to the main lesson affected them and why? (*participant responsiveness*)
5. What factors influence observed outcomes, and in what ways? (*moderators*)
  - 5.1 What school environment characteristics influence outcomes, and in what ways?
  - 5.2 What teacher characteristics influence outcomes, and in what ways?
  - 5.3 What type of pupils do teachers select for pre-teaching? What pupil characteristics influence outcomes, and in what ways?
  - 5.4 What characteristics of the pre-teaching guidance influence outcomes, and in what ways?
6. What are the causal links underpinning observed outcomes? (*mediators*)
  - 6.1 How do these work, particularly for pupils from disadvantaged backgrounds and those with lower attainment (at baseline)?
7. How well do teachers understand and comply with the outcome test administration guidelines?

## Ethics and trial registration

### Trial registration

The trial was registered on [Open Science Framework \(OSF\)](#). It was registered by the BIT Evaluation Manager on 28 August 2024. The evaluation manager will be responsible for updating the registry with the trial results after the report is published in 2026. The trial registry DOI is: <https://doi.org/10.17605/OSF.IO/UW7JB>.

### Ethical approval

Ethical approval was granted by BIT's internal ethics review committee on 08 January 2024. The review was conducted by a non-project member of staff, who blind assessed the research plan and data collection materials to ensure the project has considered all the potential risks and has appropriate procedures in place to mitigate these. In this case, the Ethical Review was conducted by Cameron Knott, Senior Adviser at BIT.

Following BIT's ethical framework, the project was assigned a medium risk level. This was due to the following key considerations:

- **Low risk.** In terms of research methods, legal exposure, and known unknowns—we used standard research methods, the legal framework in which we are operating is clear, and we had worked on similar projects before.
- **Medium risk.** In terms of participants (pupils in a school setting) and subject matter (while educational programmes may be somewhat politically contentious, this Teacher Choice does not constitute a radical change with respect to schools' usual practices).
- **High risk.** In terms of nature of data being collected (individual-level pupil data routinely collected from schools).

### Obtaining agreement to participate in the trial

We obtained agreement from schools to participate in the trial through a phased recruitment process. This included the following steps:

- Schools that expressed interest in the trial were sent an information sheet detailing the choice's objectives, implementation, and evaluation.

- This was followed up by a one to one call with the main school contact (usually the maths lead) during which we walked them through the requirements of the trial and eligibility criteria. We also explained to schools the data they would be expected to share with BIT and for what purpose.
- Schools who wished to take part were required to sign an MoU, which lays out each party's role and responsibilities to ensure the project's success. We required that the MoU was signed by a member of the school's SLT.

Schools participating in the trial also received an information sheet and withdrawal form to be shared with parents/carers of the pupils selected by teachers for the pre-teaching. These sheets were sent only to the parents/carers of the 12 pupils in each participating class identified by teachers as eligible for the pre-teaching. Any parents/carers who did not want their children's data being shared with BIT as part of the evaluation were able to withdraw their child by completing a withdrawal form and returning it to the teacher, who was asked to inform BIT (and not share their data with us). They were shared by the schools/teachers, not by BIT. The form presented information about the project and informed parents that if their child was to take part, they: i) would have to complete three short assessments (two at endline and one at baseline); and ii) may partake in a group interview with other pupils and the teacher present. The information sheets and withdrawal form (see Appendix D) contained a link to the privacy notice for the project, which can be found [here](#). We did not provide parents with the option to withdraw their child from pre-teaching (as by definition of this 'Teacher Choice', this is a teaching practice that falls under teachers and schools' discretion, and which parents would not typically be asked for their approval or be informed of).

If parents/caregivers did not opt out by returning the form, their child was included in the initial sample for the study, with schools sharing their data with an expectation that they would sit the baseline assessment. If parents/carers wished to opt out of their child's data being shared as part of the evaluation, they were asked to return the slip to teachers, who could then inform BIT (and not share that child's data).

If a parent/carer returned the form for their child (i.e. seeking withdrawal) early within the initial window (i.e. ten or more days before the end of the Summer Term), we asked teachers to identify another pupil to replace them and share the information sheet with their parent/carer in order to maximise sample size. In *this* case, withdrawal meant withdrawal from the trial. If any parent/carer returned the form after this window and/or after randomisation, we were unable to perform any 'backfilling', and withdrawals were counted toward attrition.

## Data protection

All data has been processed in accordance with the General Data Protection Regulation (GDPR) (GDPR, 2016) and Data Protection Act 2018.

## Data security

BIT has robust approaches to protecting personal data that comply with the Data Protection Act 2018 and GDPR (GDPR, 2016). We conduct all projects with a privacy by design approach to protect and maintain the privacy and security of research participants' data. All staff are trained on GDPR compliance and BIT's data protection policies. BIT is registered with the UK Information Commissioner's Office (ICO) under the terms of the Data Protection Act 2018 and has a full-time data protection officer. BIT has obtained Cyber Essentials Plus certification and International Organisation for Standardisation (ISO) 27001 certification (the international standard for information security). BIT takes steps to protect participants' personal information and prevent unauthorised access, alteration, loss, or disclosure. We train those with access to personal data, provide access only on a need-to-know basis, and remove access when it's no longer needed. Devices are encrypted, confidential data is kept securely, and confidential hard copy data (including special category data) is kept in locked rooms or cabinets. We put in place procedures to deal with any suspected personal data breach, which included notifying staff and parents and any applicable regulator of a breach where we are legally required to do so. Note that at the time of writing no such data breach has occurred.

## Data rights

Separate from the process for withdrawing a child from the evaluation (which is described above under the section on 'Obtaining agreement to participate in the trial', and which involves the deletion of all personally identifiable data associated with the pupil from BIT datasets), parents/carers could exercise a pupil's remaining data rights by contacting BIT's data protection officer via email at: dpo@bi.team. It should be noted that the extent to which these rights apply to research may vary and may be restricted in some circumstances. Typically, there is no fee required to access personal data or exercise rights, but a reasonable fee may be charged if the request is unfounded, excessive, or repetitive. If the request is deemed as such, BIT may refuse to comply with the request. BIT may request specific information to confirm the identity of the requestor and ensure that the personal data is not disclosed to an unauthorised party. In some cases, additional information may be requested to expedite the response.

BIT aims to respond to legitimate requests within a month, but if the request is complex or multiple requests have been made, it may take longer. In such a scenario, the requestor will be notified and updated on the status. It should also be noted that BIT can only comply with requests to exercise rights for personal information that directly identifies the requestor. If the information is pseudonymised or has been irreversibly anonymised and is part of the research data set, BIT will not be able to comply with the request.

## Data collected

The following data was collected from research participants to complete the trial.

### Pupil data

We collected personal data from pupils involved in the research including:

- full name;
- Unique Pupil Number (UPN);
- date of birth;
- gender;
- FSM-eligibility status;
- whether the pupil speaks English as an Additional Language (EAL);
- year group;
- class name (if more than one per year group);
- observation fieldwork notes;\*
- school name and Unique Reference Number (URN);
- results from a short assessment at the start and end of the trial (QRT);
- results from a short questionnaire about maths anxiety at the end of the trial (CMAS-UK);
- focus group discussion data—qualitative data gathered in focus groups on experiences of the pre-teaching sessions;\*\* and
- observation fieldwork notes.\*\*

\*\*This data was only ever collected for pupils in schools who agreed to take part in case studies and observations, and was therefore, not collected for every pupil involved in the One Step Ahead project. Parents/carers were informed if their child's school and/or the pupil themselves was selected for focus groups and observations, which were voluntary opt-in activities for schools.

## Teacher data

We collected personal data from teachers in participating Key Stage 1 and Key Stage 2 classes, including:

- first name;
- last name;
- work email address;
- work telephone number;
- job title;
- number of years' teaching experience;
- specialism (if any);
- usual practice of providing targeted or additional out-of-class support for Key Stage 1 Maths;
- class size;
- attitudes and views about the pre-teaching approach and the trial; and
- attitudes and views about the guidance materials for the trial.

BIT shared participants' data (as required only) with the following organisations:

- **Secure & Confidential Documents Ltd.** An independent provider contracted to print the blank QRT and CMAS-UK assessments and distribute them to schools. They only received the personal data required to print and disseminate the test papers to schools.
- **The Reasoning First Team at Oxford University.** An independent provider contracted to mark the QRT assessments at baseline and endline. They only received fully pseudonymised pupil data.
- **Qa Research.** An independent provider contracted to mark the CMAS-UK assessments at the endline. They only received fully pseudonymised pupil data.
- At the end of the evaluation, the pupil data will be shared with the EEF and FFT Education (the EEF's data processor for their archive). All the EEF trial data is stored in the EEF data archive, held within the Office for National Statistics (ONS) Secure Research Service (SRS). The archive does not contain direct identifiers like pupil name, contact details, and date of birth, but does hold a Pupil Matching Reference (PMR). The PMR is used for further matching to the National Pupil Database (NPD) and other administrative datasets that may be required as part of subsequent research. We will not use pupil names or school names in any report arising from the research. For information on how the EEF will use and protect participants' data, please see their data protection statement regarding the EEF evaluations.

## Data processing roles

Data processing roles during the evaluation up to the point of data being deleted from all locations by the evaluator and/or delivery team:

- **BIT.** Independent data controller for all data collected as part of the evaluation.
- **Participating schools.** Controllers for their pupils' data.
- **Secure & Confidential Documents Ltd.** Sub-processor.
- **The Reasoning First Team at Oxford University.** Sub-processor.

- **Qa Research.** Sub-processor.

The EEF becomes a data controller for the datasets archived after the trial, once internal quality checks have been successfully completed by the archive manager.

## Legal basis for processing data

For all information collected, BIT is relying on the lawful basis of Legitimate Interest.

BIT's lawful basis for processing personal data is based on legitimate interests according to Article 6 (1) (f) of the GDPR (2016). The processing is necessary for conducting an evaluation of the One Step Ahead Teacher Choice commissioned by the EEF, which aligns with BIT's business aims of delivering social impact through research and evaluation.

The processing of personal data from pupils, school staff, and parents is required to understand the effects of the programme, as well as to arrange observations and interviews. The processing of interview/focus group data cannot be made anonymous due to unavoidable information that may be revealed during the interviews.

The processing of pupil demographic data, participation data, and outcome data is necessary to ensure correct assessments and to match data for analysis. Anonymous data cannot be used for this purpose.

### Data retention

We only retain staff and pupil personal data for as long as necessary to fulfil the purposes we collected it for. The approximate date of deletion for staff and pupil personal data is 30 June 2026 (within six months of the final report for the One Step Ahead project being published).

## Project team

School recruitment and data collection was performed by the following BIT staff:

- Dr Laure Bokobza, Project Manager (during project set-up and implementation).
- Niall Daly, School Recruitment Manager.
- Andrés Cueto, School Recruitment Support and IPE Support.
- Tess Moseley-Roberts, School Recruitment Support.
- Lily Margaroli, School Recruitment and IPE Support.
- Hannah Bellier, School Recruitment Support and Teacher Guide Design.
- Yini Zheng, Teacher Guide Design.
- Dr Patrick Taylor, Project Director.

The evaluation was designed and delivered by the following BIT staff:

- Dr Laure Bokobza, Project Manager (during project set-up and implementation).
- Niall Daly, Project Manager (during project reporting).
- Emma Forsyth, Qualitative Research Lead.
- Trisha Harjani, Quantitative Research Lead.

- Andrés Cueto, School Recruitment and IPE Support.
- Lily Margaroli, School Recruitment and IPE Support.
- Dr Patrick Taylor, Project Director.

The following BIT researchers provided feedback and quality assurance at various points:

- Dr Giulia Tagliaferri.
- Rosario Romanos.

The following two experts supported BIT to develop the Teacher Guide and to conduct the initial pilot delivery:

- Dr Kirstin Mulholland (see biography above in the section on ‘Teacher Choice approaches’).
- Lisa Heatherington (see biography above in the section on ‘Teacher Choice approaches’).

## Methods

### Trial design

Table 5: Trial design

Trial design, including number of arms		Two-arm randomised controlled trial
Unit of randomisation		Pupil
Stratification variable (s) (if applicable)		Not applicable
Primary outcome	Variable	Mathematical reasoning skills
	Measure (instrument, scale, source)	Quantitative Reasoning Test, developed by Nunes <i>et al.</i> (2015) 20 items: sum to a maximum of 23 marks for Year 1 and 21 marks for Year 2. <i>Note the outcome is standardised across the sample</i>
Secondary outcome(s)	Variable(s)	Anxiety related to maths
	Measure(s) (instrument, scale, source)	Children Maths Anxiety Scale, developed by Petronzi <i>et al.</i> (2019) 19 items, overall score 0 to 19
Baseline for primary outcome	Variable	Mathematical reasoning skills
	Measure (instrument, scale, source)	Quantitative Reasoning Test, developed by Nunes <i>et al.</i> (2015) 20 items, overall score 0 to 20
Baseline for secondary outcome(s)	Variable	Not measured
	Measure (instrument, scale, source)	Not applicable

This trial was a two-arm RCT, with randomisation at the pupil level. An RCT design was chosen as it is considered the most robust way to evaluate the impact of pre-teaching on pupil outcomes. The two arms reflect the choice that teachers make when it comes to providing targeted support in Key Stage 1 maths, as well as being the appropriate number of arms (vs three or more) to conduct an adequately-powered evaluation with the available sample size. The two arms were:

1. **Treatment arm.** Pupils received regular targeted pre-teaching *before* the maths lessons, in fixed groups of six pupils (per class in participating year group in each school), by the teacher or teaching assistant. Each week, pupils received at least two pre-teaching sessions lasting approximately 10–15 minutes, focusing on concepts and skills in the upcoming maths lesson(s). Aside from the pre-teaching, treatment arm pupils continued to receive the in-class and out-of-class support they would usually receive in maths.
2. **Comparison arm.** Business as usual with no pre-teaching, in which the six pupils did not receive any additional support in maths *before the lesson* and otherwise continue to receive the in-class and out-of-class support they would usually receive in maths. The core restriction was that comparison group pupils do not receive maths support *before* the main lesson.

Each teacher identified 12 pupils in their class as eligible to receive the pre-teaching. This means 12 pupils in each participating Year 1 class and 12 pupils in each participating Year 2 class, in each participating school, so that the total number of 'eligible' pupils per school depended both on the number of form entry in the school and the number of available classes that school management agreed to have participate. In the spirit of 'Teacher Choices', we provided no strict guidance on pupil 'eligibility' for teachers, other than suggesting that the 12 pupils put forward in each participating class were 'likely to benefit from pre-teaching in maths', and leaving it up to teachers to apply this to their own pupils. As we required schools to provide a list of pupil names for each participating class before or at the beginning of the school year in September 2024, we encouraged teachers where possible to discuss pupil selection with the teacher of the relevant class for the previous (i.e. 2023/2024) school year. As is addressed later in the report, this broad, non-prescriptive selection method has the potential consequence of mixed-ability groups in some schools, which could dilute the pre-teaching's anticipated impact on outcomes. This factor is partially addressed by pupil-level randomisation to maximise statistical power.

We randomly allocated half of each group of 12 pupils to the pre-teaching group, and the other half to the comparison group. We chose to randomise at the pupil level as it yields the highest statistical power compared to the assessed alternative options of school, teacher, or year-group level randomisation. We estimated the risk of spillovers with pupil-level randomisation to be moderate.<sup>3</sup> Spillovers in this setting would most likely imply that the outcomes of pupils in the comparison group (not receiving the pre-teaching) are indirectly affected by the treatment given to pupils assigned to the pre-teaching group. This can be the case if class dynamics change as a result of treatment assignment in a way that affects business as usual practices, including for pupils in the comparison group, for example, if teachers change their overall pedagogical approach. It could also be a result of teachers providing more support to pupils who were allocated to the comparison group to compensate for the fact that they are not receiving the pre-teaching.

We rated this risk as moderate as our scoping research suggests that teachers were unlikely to drastically change their practices as a result of being given guidance to implement pre-teaching, and are under such time and resource constraints that their ability to provide support beyond business as usual to pupils in the comparison group should generally be limited. For spillovers to result in a dilution of the treatment effect, there would need to be a large change in the outcomes of comparison group pupils as a result of random assignment, which we did not expect.

In any case, we took the following measures to mitigate the risk of spillovers:

1. Highlighted to teachers (at recruitment and in the guidance) the importance of keeping the groups fixed for the duration of the trial.
2. Emphasised (during school recruitment and in the Teacher Guide) that pupils in the comparison group should continue to receive the in-class and out-of-class support they would otherwise typically receive, with the exception of pre-teaching, and that pupils in the treatment group should continue to receive the usual in-class and out-of-class support they would otherwise typically receive, in addition to the pre-teaching.
3. Used some of the pulse surveys during the trial to monitor compliance (we expected this to also act as a reminder of the implementation requirements for teachers).
4. Used the endline survey to measure compliance at the teacher level.
5. Used qualitative research instruments to qualify the extent and consequences of spillover effects for pupils.

Teachers delivered the pre-teaching sessions during the Autumn Term of the school year 2024/2025, with some extension into the Spring Term for those schools who began implementing the pre-teaching sessions later than planned due to logistical challenges at the start of term. The impact evaluation estimated the effect of this practice on Key Stage 1 pupils' mathematical reasoning skills and anxiety at the end of the Autumn Term. A significant minority of schools (45%) conducted their endline assessments for the trial during the Spring Term of the school year 2024/2025, due to delays in beginning the

---

<sup>3</sup> We were unable to adjust the delivery model of pre-teaching to fully ensure against the risk of spillovers/contamination here. The alternative would have been a higher level of randomisation, which would have considerably hurt the statistical power. This design decision was made following discussions with the EEF and the Study Advisory Board.

pre-teaching sessions in the Autumn Term, or for instance, a lack of resources to conduct the endline assessment in the period immediately before the festive break.

The 12 eligible pupils in each participating class completed a baseline and the endline assessment of the primary outcome measure (QRT), which captured pupils' mathematical reasoning skills. The same pupils also completed the CMAS-UK at endline only, to measure the secondary outcome of pupils' anxiety related to maths. We choose not to measure pupil anxiety at baseline to alleviate the burden on teachers and maximise cost-effectiveness.

The trial can be considered as a multi-site trial, with randomisation of pupils within schools. Given the specific type of trial—closer to an efficacy trial than an effectiveness trial—and the fact that we did not expect the sample of participating schools to be representative of the general population of primary schools in England, we aimed only at making 'conditional inference' claims. By this, we mean that the findings should be interpreted as being relevant specifically to the schools and pupils who participated in this trial (the study's 'condition'), somewhat limiting the generalisability of the results to the wider population of schools in England.

It is worth noting here that the trial design changed during implementation of the Teacher Choice, in that some schools were afforded more time than anticipated to complete delivery of the approach and completion of endline surveys. We initially planned for schools to complete the baseline assessments within the first three weeks of the Autumn Term. However, this was challenging for some schools due to the time required to administer the assessments and teacher workload being particularly high at that point of the school year. We supported schools who were struggling by granting them more time to administer the tests in situations where they were unable to meet the initial deadline to administer the baseline tests. This extension was discussed with the EEF, and aimed at retaining schools in the trial and ensuring a consistent implementation duration (and therefore pupil dosage) across schools. This change was reflected in an [updated study protocol](#) published by the EEF in December 2024 (Bellier *et al.*, 2024). Schools who started to deliver the pre-teaching sessions after 30 September 2024, based on our information, were granted an extension. They were asked to deliver the endline assessments at the end of January 2025 and to continue to deliver the pre-teaching sessions until they had done so. Overall, a large minority (45%) of schools were unable to start delivering the pre-teaching sessions before the end of September 2024. To ensure a consistent duration of implementation (and therefore dosage at the pupil level), these schools were granted an extension from end-December 2024 to end-January 2025 to administer endline assessments, and to continue to deliver the pre-teaching sessions until they had done so. We define, in our analysis in the following sections, a participating school as being 'on time' if it had started delivering pre-teaching sessions before end-September 2024, and 'late' otherwise.

## Participant selection

### School recruitment and eligibility

School recruitment took place from January to May 2024, with five additional schools recruited in July 2024 to replace schools that had dropped out of the study by that point. Mainstream, state-funded primary schools with at least one-form entry were eligible to take part in this trial (i.e. schools could not enrol pupils from classes with a mix of Year 1 and Year 2 pupils). In multi-form entry schools, as many Year 1 and Year 2 classes as the school could accommodate were permitted to take part. We aimed to recruit a minimum of 30% FSM-eligible pupils. During recruitment, our team collected data on the proportion of FSM pupils in a given school based on available government data. However, this was not a strict requirement for school eligibility. Schools were recruited nationally from all regions in England, with no regional restriction or quotas. As part of the recruitment process, eligible schools who expressed interest in the study were invited to take part in a 30-minute video call with a member of the BIT school recruitment team, who explained the project's aims, requirements, and incentives (a £300 voucher for each participating school). Following a call with schools, a member of the BIT team sent an MoU with the school for a member of its SLT to sign if they wished to enrol the school in the study.

At the end of the school recruitment process, 100 schools had signed an MoU to take part in the trial, with 29 schools withdrawing, either before baseline data collection (26 of the 29), or between baseline and endline data collection (three of

the 29). As such, pupils in 75 schools were randomised, while 71 schools participated in the full trial from baseline to endline data collection.<sup>4</sup>

### **Pupil eligibility and nomination or the trial**

In participating schools, the teachers from each Key Stage 1 year group (Year 1 and Year 2) were asked to nominate 12 pupils eligible to receive maths pre-teaching in the Autumn Term of 2024/2025. This means there were 12 pupils per class per participating year group (i.e. Year 1 and Year 2). In a single-form entry school, we therefore had a total of 24 pupils participating in the trial, combining the treatment and comparison arms. In a two-form entry school, there were 24 (12 + 12) in Year 1 and 24 in Year 2, so 48 pupils nominated for participation in the trial in total.

We asked teachers to identify and nominate eligible pupils by the end of June 2024. We sent written guidance to teachers in participating schools via email, to remind them that they should seek to select pupils whom they think would benefit most from this additional support. This could be pupils who were working below age-related expectations, experience maths anxiety, struggle with some key concepts and skills, or have additional learning needs for example. The specific language we communicated to teachers was that they should identify and put forward 12 pupils in each participating class who ‘could benefit from additional support and more active engagement in maths’. This guidance formed part of the Teacher Guide we shared with schools after they signed up for the study, and which we include here in [Appendix C](#).

There were no strict inclusion or exclusion criteria for pupils in the pre-teaching group, and we intentionally left the choice to teachers’ discretion, emphasising that ‘they know their pupils best’. Given that the majority of teachers did not teach the pupils during the previous (2023/2024) academic year, we invited teachers to discuss pupil nomination with the previous year teacher (i.e. the Year 2 teacher nominating incoming Year 2 pupils was invited to discuss pupil nomination with the Year 1 teacher who had taught the pupils during the previous academic year, and the same for Year 1 teachers with Reception teachers for the previous academic year). Findings from the scoping research and Teacher Guide pilot suggest that teachers were likely to use a mix of objective criteria (e.g. assessment scores) and subjective ones (e.g. pupil attitudes and behaviour during the maths lesson). Specifically, out of the six teachers who took part in the Teacher Guide pilot, they reported using a range of criteria for selecting pupils for pre-teaching: previous attainment and age-related expectations were selected to the same extent as attitudes towards maths and engagement in the maths lesson. We aimed to probe the factors driving teachers’ selection of pupils in the baseline survey of teachers.

Nominated pupils were then included in the trial on an opt-out basis, with schools asked to share a parent/carers information and withdrawal form with the parents/carers of nominated pupils (included here in [Appendix D](#)). If a nominated pupil’s parents/carers did not withdraw them from the study by a stated date (which changed depending on when we asked the school to send the form out), the pupil was included in the list of those for whom a baseline assessment was prepared (see ‘Baseline and endline surveys’ subsection under ‘Research methods’ section below). Pupils who completed a baseline assessment were then randomised into either the treatment (allocated to receive pre-teaching for Key Stage 1 maths) or control (business as usual support) conditions. At the end of the pupil enrolment recruitment process, 2,609 pupils had completed the baseline QRT assessment and been randomised into one of the two conditions.

Note that pupils could still withdraw their participation from the evaluation at any point, even after completion of a baseline survey (and being randomised into the treatment or control conditions). In total, 262 pupils withdrew after having been randomised into one of the treatment conditions (i.e. after completion of a baseline survey). Of this 262, 164 were withdrawn by their school on a whole class/year basis (i.e. the school stating that a particular class or year group would no longer be able to participate in the trial),<sup>5</sup> with 98 withdrawn individually by parents/carers at some point over the course of the study.

---

<sup>4</sup> One school received its randomisation allocation after it had informed BIT that it had completed administering baseline assessments, however, the school then withdrew from the study after receiving its randomisation allocations, without ever having shared the baseline data with BIT. We consider this an unusual case to be both ‘randomised’ and ‘withdrawn before baseline data collection’.

<sup>5</sup> These 164 pupils represent 14 classes withdrawn on a ‘whole class/year’ basis across seven schools.

## Outcome measures


### Baseline measures

We used the Quantitative Reasoning Test (QRT) to measure pupils' baseline mathematical reasoning skills. This was the sole baseline measure. Please see the subsequent section 'Primary outcome' below for more details on the QRT.

### Primary outcome

The primary outcome was pupils' mathematical reasoning skills. We measured this using the QRT at baseline and endline for all pupils nominated as eligible to receive the pre-teaching irrespective of their treatment assignment. We worked with the assessment's developer, Dr Terezinha Nunes, and the Reasoning First team to tailor the assessment to the trial's aims. In order to monitor any reasoning progress, the team curated two versions of the assessment for each year group; a baseline test and a more challenging endline test. Each test contained 20 items from a wider pool of items and consisted of four subsections covering arithmetic, additive composition, inverse relation, and multiplicative reasoning in simple contexts. These four subsections were selected because they represent the foundational cognitive components of mathematical reasoning required at Key Stage 1, as identified by the assessment developers. In this scale, each 'item' refers to a reasoning problem. The items were selected by Dr Terezinha Nunes and her team, for each year group respectively, building from the findings of previous studies using the test. The resulting test has five items per scale, with a total of 20 items. The structure of the items was the same at baseline and endline, given the relatively short time period in between. The text of the corresponding questions, however, was different. Below we reproduce a question that was asked with two versions at baseline and endline. The baseline and endline tests were intentionally designed to differ in difficulty. Given the rapid developmental progression of Key Stage 1 pupils, a test appropriate for September would potentially result in ceiling effects by December/January, which would have masked potential effects. Therefore, the test developers created an endline assessment with higher difficulty to accurately capture the full range of attainment at that later time point. To ensure valid comparison despite these differences in raw difficulty, all scores were standardised within the analysis (as detailed in the 'Statistical analysis' section).

Table 6: Extracts from the QRT at baseline and endline

Baseline	Endline
<p><i>Item 8</i></p> <p>THE SAME COINS ARE STILL IN FRONT OF THE CHILD.</p> <ul style="list-style-type: none"> <li>8a: Another day it was your turn to go to the shop. You wanted to buy a rubber. This is the money you had. The rubber costs 9p. You want to pay the exact money. Take the coins that you need to pay exactly 9p. Record the amount of money that the child took or write zero, if the child said that there was not enough money.</li> <li>8b: If the child says there is not enough money or counts the 5p coin as 1p, ask: <b>How much money is this?</b> Show the 5p coin. If the child does not say 5p say: <b>This is 5p. You need to say five when you point to the 5p coin. Can you check if you have enough money?</b> Record the amount of money that the child took or write zero, if the child said that there was not enough money.</li> </ul> <p>THIS IS A SECOND TRIAL AND YOU RECORD AS 8B. IF YOU DID NOT USE 8B, PUT AN X WHERE THE ANSWER WOULD GO.</p> <p>THERE IS NO SECOND TRIAL FOR THE OTHER MONEY ITEMS.</p> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin-left: auto; margin-right: auto;"> <p style="text-align: center; margin: 0;"><b>Record</b></p> <p>8a: money paid for rubber: _____</p> <p>8b: money paid for rubber: _____</p> </div>	<p><i>Item 7</i></p> <p>REMOVE THE 2p COINS AND PLACE ONE 5p COIN AND TEN 1p COINS IN FRONT OF THE CHILD.</p> <ul style="list-style-type: none"> <li>7a. On another day Grace was the one who went to the shop. Grace wanted to buy a rubber. This is the money she had. SHOW ONE 5p COIN AND ASK. Do you know how much money this is?</li> <li>IF THE CHILD DOES NOT KNOW, YOU SAY: <b>This is 5p.</b> ALL THE CHILDREN HAVE A REMINDER: SHOW ONE COIN OF EACH VALUE AND REMIND THE CHILD: <b>This is 1p and this is 5p.</b></li> <li><b>This is the rubber Grace wanted to buy. The rubber costs 9p. Grace wanted to pay the exact money. Take the coins that Grace needed to pay exactly 9p.</b></li> </ul> <p>RECORD THE AMOUNT OF MONEY THAT THE CHILD TOOK. IF THE CHILD SAID THAT THERE WAS NOT ENOUGH MONEY, WRITE ZERO.</p> <p>7b: IF THE CHILD SAYS THERE IS NOT ENOUGH MONEY, REMIND THE CHILD: <b>This is 1p and this is 5p. When you point to the 5p coin, you need to say five. Can you check if you have enough money?</b> RECORD THE AMOUNT OF MONEY THAT THE CHILD TOOK OR WRITE ZERO, IF THE CHILD SAID THAT THERE WAS NOT ENOUGH MONEY. IF YOU DID NOT USE 7B, PUT AN X WHERE THE ANSWER WOULD GO.</p> <p>THERE IS NO SECOND TRIAL FOR THE OTHER MONEY ITEMS.</p> <div style="text-align: center; margin: 10px 0;">  </div> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin-left: auto; margin-right: auto;"> <p style="text-align: center; margin: 0;"><b>Write on the record sheet:</b></p> <p>7a: Money paid for the rubber: _____</p> <p>7b: Money paid for the rubber: _____</p> </div>

We consulted Dr Terezinha Nunes to ensure the test was well-suited to capture mathematical reasoning skills for pupils in Years 1 and 2, and to obtain insights on administration modalities. They recommended administering the test in a one to one format for Year 1 pupils, and in a whole-group format for Year 2 pupils. It was estimated that the one to one tests would take approximately 30 minutes to complete per pupil. In contrast, for Year 2 assessments, it was anticipated that 45 to 60 minutes was sufficient time for the assessment delivery to the whole group (12 pupils per class). We provided teachers with guidance and resources to help in administering the QRT, which are included in [Appendix E](#). The Reasoning First team developed training materials for Year 1 and Year 2 teachers to support them to administer the assessment. Due to the constraints of the overall evaluation timeline, these materials were not piloted prior to use in the trial. Teachers were asked to watch a short online training video (30 minutes) at their convenience, prior to baseline administration. We also shared short written guidance summarising and complementing the training, for them to refer back to.

For the Year 1 test, teachers were required to use materials commonly available in classrooms (coins, blocks or tokens, small dolls) and needed printed paper copies of the testing instructions, of the detailed item-by-item instructions, and of one record sheet per pupil. For the Year 2 test, teachers used a screen in front of the class to project PowerPoint slides, printed item-by-item instructions for test delivery and printed answer booklets, one per pupil. The Reasoning First Team at Oxford University marked the baseline tests from December 2024 to March 2025, and the endline tests from March to June 2025.

Having teachers administer the test themselves had both some advantages and drawbacks. It increased the burden on teachers, as they were required to spend time familiarising themselves with the assessment and delivering it to pupils. It also introduced a risk of bias in how answers were recorded, given that teachers were not blind to pupil assignment to the trial arms. The alternative would have been to hire an external party to send assessors into schools to deliver the assessments. Teacher-led administration also presents some advantages, including being in line with the broader aims of Teacher Choices research of empowering teachers to implement and test new approaches in their classrooms. We expected teacher-led administration to be beneficial for pupils, particularly for the younger age group (Year 1 pupils) for whom the assessment was administered on a one to one basis. Having a trusted adult deliver the assessment was likely to put pupils more at ease and boost their confidence. The Year 1 version of the assessment also entailed recording pupils' behavioural responses. Dr Terezinha Nunes advised that qualified class teachers and teaching assistants were better placed to deliver this assessment than trained external assessors. We also took into consideration that having teachers administer the test was also the most cost-effective option.

To maintain consistency in measurement, we decided on teacher administration for both year groups. To reduce the burden on teachers, we ensured the training for teachers was as flexible and as short as possible, and available in a mix of formats (online videos and written guidance). The training materials were developed with the specific aims of: i) improving the quality and consistency of assessment administration; and ii) reducing any bias arising from teacher administration of the tests. External marking was implemented to reduce the risk of bias, with markers blinded to treatment group assignments. To ensure quality and consistency across markers, a 20% sample of tests was independently double-marked. This quality control measure mitigated potential human error and subjective interpretation of answers. As part of these quality control measures, the test developers collaborated directly with the research assistants to standardise the marking process.

For Year 1 tests, the quality control procedure involved research assistants making a note of anything unusual on a test or batch of tests, so that all tests from the relevant school could be reviewed. Examples of such cases included instances where answers appeared to be entered by a pupil rather than the teacher (as should have been the case for Year 1 tests), incomplete tests whose administration seemed to have been interrupted with no record of answers after a certain item, or the teacher's entry for a question part being inconsistent with the pupil's final answer. Furthermore, if a pupils' writing was difficult to interpret, such as mirror images of a number, the test was marked for a second time by a different research assistant. A score of 0 was assigned to pupils under two scenarios: i) if the pupil was given the test but answered incorrectly; or ii) if the test was blank but the pupil was confirmed as being present for the administration of the assessment. The quality assurance process began with overall checks on all tests before samples were selected for double-marking, prioritising schools where potential issues were identified.

To ensure consistency and comparability of results, the QRT scores were standardised for analysis. This is a deviation from the study plan (Bellier *et al.*, 2024), however, it was deemed necessary as the QRT tests administered to Year 1 and Year 2 pupils had different score ranges (e.g. Year 1 baseline out of 23 and endline out of 24, while Year 2 was out of 21 for both baseline and endline).<sup>6</sup> The endline QRT test for each year group was also designed to be more challenging than its baseline counterpart.

To standardise, we rescaled pupil scores within each year group (Year 1 and Year 2). For each group, we calculated the mean and the pooled standard deviation (SD) of scores across both treatment and comparison group pupils. Each individual's score was then expressed relative to this distribution by subtracting the group mean and dividing by the group SD. The resulting values represent how many SDs above or below the average a pupil scored, which allows us to compare outcomes across groups despite the slightly different test score ranges. This standardisation aligns with the EEF guidance, which permits data transformation, such as using a standardised score, when there are 'multiple cohorts using different pre-tests or varying levels of difficulty' (EEF, 2022). Standardising the scores allows for meaningful comparisons of pupil mathematical reasoning skills across these different year groups and test forms. The standardisation was performed by using the pooled SD across treatment and comparison groups for each respective year group.

### Secondary outcome

The CMAS-UK served as the secondary outcome measure aimed at assessing pupils' anxiety related to mathematics. At the time of planning, we consulted the developers of the scale, Dr Dominic Petronzi and Dr Thomas Hunt, to ensure its suitability for the study's aims. They confirmed that the scale was developed and validated for pupils aged 4 to 7 in the UK. The scale has historically exhibited high internal consistency, with a Cronbach's Alpha reliability coefficient of 0.88 (Petronzi *et al.*, 2019). The CMAS-UK comprises 19 items, and teachers were provided with a separate set of instructions on how to administer the CMAS-UK.

We also consulted the test developers, Dr Dominic Petronzi and Dr Thomas Hunt, to provide insights into administration modalities. They advised that Year 1 and Year 2 pupils can complete the scale in a group setting with an ideal maximum of eight pupils and with a teacher or teaching assistant setting the pace, typically taking 30 minutes. We recommended providing a space (e.g. an empty) chair between pupils when they complete the assessment, to avoid pupils copying each other. For each item, teachers read out statements, and pupils respond by circling one of three faces—a happy face, a neutral face, or a sad face—to indicate their feelings in various typical maths lesson scenarios. These statements cover a range of situations, such as a pupil's feelings when entering a numeracy class, when peers finish work quickly, or when they make a mistake in numeracy. While the items are designed to capture underlying concepts like low self-efficacy or the experience of failure, the scale is intended to measure overall pupil anxiety, and is not structured to capture sub-components separately.

The regression analysis used each pupil's average score across all questions they attempted, meaning that skipped (i.e. unattempted) questions were excluded from the calculation. This approach ensured comparability across pupils, as the average score is not biased by differences in the number of questions attempted (e.g. a pupil attempting ten questions vs one attempting 19). The tests were administered by teachers, but were marked externally by Qa Research, with the external markers being blind to treatment condition assignment to minimise the risk of bias.

### Sample size

At the design stage, the study aimed to recruit 95 schools and 2,280 pupils. This initial calculation assumed 95 single-form entry schools, with 1,140 pupils in the pre-teaching group and 1,140 in the comparison group. However, after recruitment and accounting for multi-form entry schools, the final estimated sample size (before school and/or pupil withdrawals) was 3,708 pupils across 95 schools, implying an expected 1,854 pupils in the pre-teaching group and 1,854 in the comparison

---

<sup>6</sup> The [original study plan](#) stated the QRT outcome measure would be the raw overall score (0 to 20), but this approach was abandoned as the tests used varying score ranges for different year groups and time points (Bellier *et al.*, 2024).

group.<sup>7</sup> The assumed sample size at the design stage (2,280 pupils) implied that we were powered to detect a minimum detectable effect size (**MDES**) of **0.125** for the primary outcome analysis and **0.23** for the FSM analysis. The realised sample size post-recruitment (but before school and pupil withdrawal) implied equivalent **MDES figures of 0.098** and **0.23**, respectively. It was assumed at the design stage that 684 out of the anticipated 2,280 pupils (30%) would be FSM-eligible, and the actual figure in the realised recruited sample (before school and pupil withdrawal) had 1,112 of 3,708 pupils (29.99%) as FSM-eligible.

Key assumptions for the power calculations included:

- An approximate 6% school dropout rate and 14% pupil dropout rate, resulting in an overall 19.16% dropout.
- A pre-/post-test correlation of 0.29, derived from the EEF's '1stClass@Number 1' trial (Sahasranaman *et al.*, 2025), which used the same primary outcome measure.
- No intracluster correlation coefficients (ICCs) were assumed at the class or school level, as randomisation occurred at the pupil level.

The choice of MDES was influenced by existing evidence on the impact of pre-teaching in maths for this age group, which was limited. However, all available literature suggested an effect size between 0.10 and 0.25. All significance tests were two-sided, with a 5% significance level.

## Randomisation

The study used randomisation at the pupil level within participating schools and classes. From the 12 pupils nominated by each class teacher as eligible for pre-teaching, six were randomly assigned to the treatment (pre-teaching) group, and the other six were allocated to the comparison (no pre-teaching, business as usual) group.

The randomisation process was carried out by BIT using anonymised pupil-level data. The sequence was generated in R by a member of the BIT project team and was internally quality assured by an independent researcher outside the project team. After randomisation, BIT sent schools the lists of treatment and comparison pupils.

We opted for simple pupil-level randomisation without stratification for practical and theoretical considerations. Although BIT collected pupil data for eligible participants, complete data was not always available before the trial began. Stratifying class-by-class with a maximum of 12 eligible pupils per class was deemed challenging, as it could result in very small strata, potentially leading to imbalanced groups or 'over-stratification'. Instead, relevant variables were accounted for in an additional sensitivity analysis. Specifically, in order to account for potential imbalances, the 'Additional analyses and robustness checks' section reports a model that includes the following covariates: year group; gender; FSM status; and EAL. The primary risk associated with avoiding stratification was a potential imbalance between the treatment and comparison groups on key pupil characteristics (such as baseline QRT score, FSM status, gender, or year group), which could threaten the internal validity of our main estimate if unadjusted. This risk was systematically mitigated in the analysis by applying a regression-based adjustment that included the baseline QRT score as a covariate, and further guarded against by testing the balance of all key demographic characteristics (year group, gender, FSM status, and EAL status) as part of overall robustness checks.

The randomisation took place for each school only after the school project lead confirmed the completion of all baseline assessments. This timing was specifically chosen to ensure that teachers remained blind to pupils' random assignment during baseline test administration, thereby mitigating potential bias in the baseline test data. Once BIT received confirmation that baseline assessments had been completed, participating pupils in a school were randomised into either the pre-teaching or control conditions, with this allocation securely shared with each school on a rolling basis. Note that to ensure that schools could begin planning for (and delivering) pre-teaching sessions as soon as possible and stick with the

---

<sup>7</sup> While 100 schools overall signed an MoU, 95 did so during the school recruitment period (January to May 2024), with an additional five schools recruited in July 2024 to replace five schools that had withdrawn soon after being recruited. As such, there were a maximum of 95 schools participating in the trial at any one time.

tight planned implementation timelines, we performed randomisation and shared the pupil allocations with each school without necessarily waiting for the baseline assessments to arrive at BIT (via secure courier service).

## Statistical analysis

Analysis was conducted on an intention-to-treat (ITT) basis, including all pupils randomised into either the pre-teaching or comparison groups. Analysis was conducted in RStudio, using two-sided significance tests, at the 5% significance level. The analysis compares the impact of regular pre-teaching on pupils' mathematical reasoning skills (treatment group), with business as usual support in maths excluding pre-teaching (comparison group).

### Primary analysis

The primary analysis included all eligible pupils in participating schools (i.e. pupils who were randomised into either the pre-teaching or comparison groups) for whom we had complete data on the relevant variables, in line with the EEF's statistical analysis guidance (EEF, 2022). Analysis was carried out using an ordinary least squares (OLS) regression with school-level fixed effects. The fixed effects model accounts for the risk of false positives that arises from the clustering of observations within schools.

$$Y_{is} = \beta_0 + \beta_1 T_{is} + \beta_2 PreQRT_{is} + \gamma_s + \epsilon_{is} \text{ (Equation 1)}$$

Where:

- $Y_{is}$  is the endline mathematical reasoning for individual  $i$  from school  $s$ , as measured by the endline QRT administered at the end of the Autumn Term.
- $T_{is}$  is a pupil-level dummy indicator of whether pupil  $i$  from school  $s$  was randomly assigned to the pre-teaching group, or to the comparison group with no pre-teaching.
- $PreQRT_{is}$  is the baseline mathematical reasoning for individual  $i$  from school  $s$ , as measured by the baseline QRT administered at the beginning of the Autumn Term.
- $\gamma_s$  represents the school-level fixed effects, capturing unobserved school-specific characteristics.
- $\epsilon_{is}$  is the heteroskedasticity-robust Huber-White error term for pupil  $i$  from school  $s$ .

### Secondary analysis

#### *Pupil anxiety related to maths*

Secondary analysis was also conducted on an intention-to-treat (ITT) basis using the specification outlined below. Analysis was carried out using an OLS regression with school-level fixed effects:

$$Y_i = \beta_0 + \beta_1 T_{is} + \gamma_s + \epsilon_{is} \text{ (Equation 2)}$$

Where:

- $Y_i$  is the individual's anxiety related to maths as measured by the CMAS-UK administered at the end of the Autumn Term.
- $T_{is}$  is a pupil-level dummy indicator of whether pupil  $i$  from school  $s$  was randomly assigned to the pre-teaching group, or to the comparison group with no pre-teaching.
- $\gamma_s$  represents the school-level fixed effects, capturing unobserved school-specific characteristics.
- $\epsilon_{is}$  is the heteroskedasticity-Huber-White error term for pupil  $i$  from school  $s$ .

## Analysis in the presence of non-compliance

In the study plan (Bellier *et al.*, 2024), we pre-specify that we will use teacher responses to the endline surveys to get a rough indicator of pupil-level compliance. It was intended that all pupils in the respondent's (teacher's) class would be classified as compliant/non-compliant depending on the teacher's responses. Specifically, we pre-specify that we will obtain a Complier Average Causal Effect (CACE) estimate by dividing the ITT estimate from the primary analysis by the proportion of compliers in the sample, with two thresholds for compliance: i) fully and partially compliant; and ii) fully compliant only.

However, due to both inconsistencies in the format and recording of class names/numbers in the data provided by schools, and consequently in order to limit the data collection burden in the surveys, teachers were not asked to identify their class. This meant that any linkage of teacher and pupil data was not possible beyond the school level. Due to this data limitation, we do not derive a CACE estimate, which is a deviation from the published study plan (Bellier *et al.*, 2024). The lack of a derived CACE estimate prevents a more sophisticated isolation of pre-teaching's true impact on only those pupils who fully received the intended dosage, and limits the interpretation of results strictly to the ITT estimate. Instead, we report levels of compliance at the teacher level descriptively and provide further breakdowns by criteria. The levels of compliance are further explored in the '[IPE](#)' and '[Conclusion](#)' sections also.

## Missing data analysis

There are three possible sources of missing data:

1. Missing teacher/teaching assistant data: non-response to baseline or endline surveys.
2. Missing pupil outcome data.
3. Missing pupil covariates (e.g. demographic data).

We first report the number of missing observations for all three cases. Missing data of type 1 is reported in the '[IPE results](#)' [section](#) below, and we do not perform any missing data robustness analysis for the teacher survey responses.

For missing outcome data (type 2) and missing pupil pre-treatment covariates (type 3), we check whether any of these constitute over 5% of the observations and whether there is a difference in the proportion missing between treatment and comparison groups. The rule pre-specified in the evaluation's study plan (Bellier *et al.*, 2024) was that if these observations were under the 5% threshold and balanced between the two groups, we would proceed with the specified analysis in RStudio (including only complete cases). Conversely, if over 5% of the observations were missing for pupil outcome data and/or pupil pre-treatment covariate data, we were committed in the study plan (Bellier *et al.*, 2024) to taking the following steps to understand whether the data were Missing at Random (MAR) or Missing Not at Random (MNAR).<sup>8</sup>

In the 'Impact evaluation results' section, we examine the first two types of missingness in line with the study plan and the EEF statistical guidance (Bellier *et al.*, 2024; EEF, 2022). Specifically, for the cases of any missing covariate (QRT baseline), and outcome (QRT endline) data over 5%, missingness patterns was examined using logistic regression models to predict missingness. Specifically, a binary variable was created where the value 1 was assigned if any of the covariate or outcome data is missing, and value 0 if it was a complete observation. This allows for the regression to predict *missingness*. The logistic regression then investigated whether this binary missingness variable was predicted by any of the covariate data. The results of this analysis are included in the '[Impact evaluation results](#)' [section](#) below.

## Missing covariate data

The sole covariate used in the primary and secondary analyses is the baseline QRT score. The study plan (Bellier *et al.*, 2024) indicates that if missingness was predicted by other observed covariates or outcomes (MAR), complete case analysis may yield biased estimates. The study plan (Bellier *et al.*, 2024) specifies that we may re-run the primary outcome analysis with multiple imputation. Note that in the realised analysis we did not proceed with multiple imputation, as the missingness of

---

<sup>8</sup> Missing at Random (MAR) occurs when the missingness in a variable can be predicted from other observed variables, while Missing Not at Random (MNAR) occurs when missingness depends on an unobserved variable even after taking into account all the information in the observed variables.

the sole covariate (QRT baseline) does not cross the 5% threshold, and missingness is not predicted by the observed covariates. We report this in more detail in the [‘Impact evaluation results’ section](#) below.

### Missing outcome data

The study plan (Bellier *et al.*, 2024) specifies that if the data are found to be Missing at Random (MAR), that is, predicted by other observed covariates, we would re-run a complete case analysis with and without covariates. Note that the observed level and type of missingness did not require us to re-run the analysis in that regard.

### Subgroup analyses

#### *FSM-eligible pupils*

We performed two subgroup analyses for FSM-eligible pupils, to study any heterogeneous treatment effects in that group, following the EEF statistical analysis guidelines (EEF, 2022):

1. Estimate the model specified in the primary analysis on the subsample of FSM-eligible pupils.
2. Estimate a similar model including an interaction term for FSM eligibility, using the entire sample (see Equation 4 below).

Both approaches can be used to estimate the effect size for FSM pupils, but the latter uses information from the whole sample. The effect size of subgroup analysis 1 and the interaction term coefficient from subgroup analysis 2 with an associated measure of uncertainty, are reported in the [‘Impact evaluation results’ section](#) below. Model 1 is reported throughout the main report as the headline estimate, however, the effect sizes of both analyses are compared and reported in [Appendix H](#), as a sensitivity and robustness check for subgroup analysis 1. In principle, while these two methods differ, they should provide analogous results. We explore any difference and discuss its implications for the results in the [‘Impact evaluation results’ section](#) below.

$$Y_{is} = \beta_0 + \beta_1 T_{is} + \beta_2 FSM_{is} + \beta_3 (FSM_{is} * T_{is}) + \beta_4 PreQRT_{is} + \gamma_s + \epsilon_{is} \text{ (Equation 4)}$$

Where:

- $Y_{is}$  is the endline QRT overall test score for individual  $i$ , in school  $s$ .
- $T_{is}$  is a pupil-level dummy indicator of whether pupil  $i$  from school  $s$  was randomly assigned to the pre-teaching group, or to the comparison group with no pre-teaching.
- $FSM_{is}$  is a binary indicator equal to 1 if the pupil is eligible for FSM in the 2024/2025 academic year and 0 if not.
- $PreQRT_{is}$  is the baseline mathematical reasoning for individual  $i$  from school  $s$ , as measured by the baseline QRT administered at the beginning of the Autumn Term.
- $\gamma_s$  represents the school-level fixed effects, capturing unobserved school-specific characteristics.
- $\epsilon_{is}$  is the cluster-robust error term, for individual  $i$  in school  $s$ , clustered at the pupil level.

We report and compare the effect sizes of both models, the restricted sample and the interaction term. The effect sizes are reported in terms of Hedges’  $g$ .

#### *‘On time’ versus ‘late’ schools*

We perform a subgroup analysis to study potential heterogeneous treatment effects for pupils in ‘on time’ schools relative to ‘late’ ones. We do this by adding an interaction term for the status of the school (‘on time’ or ‘late’) with the treatment binary variable, defined at the pupil level. This approach allows us to test directly whether the treatment effect varies by ‘on time’ versus ‘late’ schools. A large minority (45%) of schools were unable to start delivering the pre-teaching sessions before the end of September 2024. To ensure a consistent duration of implementation (and therefore dosage at the pupil level), these schools were granted an extension from end-December 2024 to end-January 2025 to administer endline assessments, and to continue to deliver the pre-teaching sessions until they had done so. We define a participating school

as being ‘on time’ if it had started delivering pre-teaching sessions before end-September 2024, and ‘late’ otherwise. See ‘Outcome measures’ section above.

There are some limitations to this specification. Because ‘on time’/‘late’ status is a school-level variable and we already include school-level fixed effects, the main effect of ‘on time’/‘late’ status will be perfectly collinear with the fixed effects and automatically will drop out of most regressions performed by statistical software. This is a feature of the model, and does not mean we are not measuring the interaction between ‘on time’/‘late’ status and treatment correctly. Furthermore, we note that school-level fixed effects already account for unobserved differences between schools, which might also absorb some variation in treatment effects between the types of schools (in other words, there may not be much within-school variation left to leverage). Lastly, the statistical power for this analysis will be lower than for the main analysis (while we are including the whole sample of schools and pupils, the expected effect size is lower for the interaction term than for the main treatment effect).

However, this was still worthwhile as an exploratory subgroup analysis, to test whether the start of implementation timing affected pupils’ outcomes.<sup>9</sup> Given that this variation was an unplanned variation in the study design, we also report these results for transparency in the ‘[Impact evaluation results](#)’ section below.

$$Y_{is} = \beta_0 + \beta_1 T_{is} + \beta_3 (OnTime_s * T_{is}) + \beta_4 PreQRT_{is} + \gamma_s + \epsilon_{is} \text{ (Equation 5)}$$

Where:

- $Y_{is}$  is the endline QRT overall test score for individual  $i$ , in school  $s$ .
- $T_{is}$  is a pupil-level dummy indicator of whether pupil  $i$  from school  $s$  was randomly assigned to the pre-teaching group, or to the comparison group with no pre-teaching.
- $OnTime_s$  is a binary indicator equal to 1 if the pupil’s school is defined as ‘on time’, and 0 if it is defined as ‘late’.
- $PreQRT_{is}$  is the baseline mathematical reasoning for individual  $i$  from school  $s$ , as measured by the baseline QRT administered at the beginning of the Autumn Term.
- $\gamma_s$  represents the school-level fixed effects, capturing unobserved school-specific characteristics.
- $\epsilon_{is}$  is the cluster-robust error term, for individual  $i$  in school  $s$ , clustered at the pupil level.

The effect sizes are reported in terms of Hedges’  $g$ .

For all the models above, all covariate data, including FSM status, were gathered at the pupil level from schools, with data provided by all schools using a clear template provided by BIT, which avoided the need for additional data cleaning. We chose this route rather than gathering covariate data from the National Pupil database (NPD), for time efficiency, with the aim of being able to disseminate results to teachers and schools as quickly as possible. Applications to access the NPD typically take at least six months and delays are frequent. Given that the outcome data was collected directly in schools, the potential for increased precision of covariate data taken from the NPD was not considered to outweigh the resulting change in the time frame.

### Additional analyses and robustness checks

As robustness check, we re-run Equations 1, 2, and 4 including in addition  $X_{is}$ , a vector of pupil and school-level covariates including year group, gender, and EAL status for individual  $i$ , in school  $s$ . We selected these covariates based on relevance to the analysis. We include EAL status as a covariate given it could account for some of the variance. Specifically, pupils at the Year 1 and Year 2 level are developing mathematical literacy in addition to mathematical logic. Hence, their language ability may have a direct impact on their mathematical reasoning.

---

<sup>9</sup> We perform this exploratory analysis based on the *ex post* hypothesis that starting the delivery of pre-teaching later also meant administering the endline assessment further into the busy school year (January), a factor which may exacerbate implementation challenges related to staff capacity, consistency of delivery, and pupil engagement.

We also estimate Equation 2 (secondary outcome) with the baseline QRT score as a robustness check.

#### *Estimation of effect sizes*

The primary and secondary analysis, including the FSM subgroup analysis, convert effect sizes into Hedges'  $g$  using the following transformation:

$$\text{Hedges' } g = \frac{M_1 - M_2}{SD^*_{pooled}} \text{ (Equation 3)}$$

Where  $M_1 - M_2$  is the difference in mean test scores between the treatment and comparison group, and  $SD^*_{pooled}$  is the pooled SD adjusted for the school fixed effects.

We report 95% confidence intervals (CIs) for the effect sizes of all primary and secondary outcome analyses.

## IPE

The IPE was designed with two key aims:

- to assess delivery of maths pre-teaching sessions for pupils allocated to the treatment group; and
- to explain findings from the impact evaluation, through exploring causal links articulated in the logic model and barriers and facilitators to delivering pre-teaching.

The full list of research questions for the IPE are set out in the 'Introduction' section. To address these questions a mixed-methods approach was used, combining quantitative methods (surveys administered to all schools participating in the trial) with qualitative research conducted in six case study schools. The methodology was informed by the desire to triangulate information from different participant groups (teaching staff and pupils) while keeping the data collection requirements for schools proportionate to the evaluation as a Teacher Choice trial (which are intended to be less burdensome on teachers).

## Research methods

### *Baseline and endline surveys*

All schools participating in the trial were invited to complete a survey at baseline (05 September 2024 to 26 September 2024) and endline (12 December 2024 to 10 March 2025).<sup>10</sup> These were conducted online using SmartSurvey, an online survey platform. The survey was sent to all teachers for whom we had contact details, so that we could capture multiple perspectives at each school (baseline, n=320, endline, n=273).

At the start of the baseline survey, respondents were asked to select their role on the project (as delivering pre-teaching, school trial lead or both). Questions covered the following topics:

- school context (class size, staffing, key challenges)—only school trial leads were asked these questions;
- knowledge and experience of pre-teaching, including confidence to deliver;
- how Key Stage 1 maths is taught—including details of business as usual targeted support; and
- how pupils were selected, and their characteristics.

---

<sup>10</sup> The surveys were kept open for longer than planned to give schools that started implementing pre-teaching later an opportunity to complete them.

To encourage responses to the baseline survey, the initial link was followed by three reminders sent on 10, 18, and 23 September 2024. The baseline survey received 108 responses (a response rate of 34%).

Respondents who completed the survey were also entered into a prize draw for a chance to win a £50 voucher at baseline, and one of two £50 vouchers at endline.

The endline survey was administered at the end of delivery to gather data on the following:

- how sessions were delivered (to capture information on compliance);
- perceptions of pupil engagement and outcomes; and
- teachers' experience of participating in the trial (including administering assessments).

Respondents who did not complete the baseline survey were also shown questions on school context, pre-teaching knowledge/experience, and usual practice to retrospectively gather this information. At endline, follow-up emails were tailored to four teacher subgroups based on school timing and baseline completion status. Two reminders were sent per subgroup: on 08 and 21 January 2025 for 'on time' schools, and on 25 February and 05 March 2025 for 'late' starting schools. The endline survey received a total of 126 responses (a response rate of 46%). It is worth noting that the relatively low response rate for both the baseline and endline survey limits the generalisability of the associated findings presented later in the report, given that they may not reflect the views of the 'least engaged' half of the sample of participating teachers.

Both surveys consisted of closed and open-text questions, the latter was kept to a minimum and used mainly where a more detailed response was needed.

#### *Pulse surveys*

Short, online 'pulse' surveys were sent to teachers who were delivering the pre-teaching at regular intervals during delivery, in all schools. In total, three pulse surveys were sent out in September, October, and November 2024 (more details are set out in Table 7 below).<sup>11</sup> Pulse survey 1 was designed to get rapid feedback on fidelity and capture initial impressions of engagement and ease of implementation. We introduced a question on challenges in pulse survey 2. In Pulse survey 3, we introduced some new topics that we felt teachers would be in a better place to respond to having done more of the delivery.

The pulse surveys were designed to be a more engaging, less burdensome alternative to other methods of collecting monitoring data. As a new type of data collection tool, it was difficult to predict how many teachers would engage. The lowest rate for Pulse survey 1 may reflect timing as this was sent out relatively early in the term when teachers had only recently begun to implement the pre-teaching, and this may have been a barrier to engagement. Response rates did improve throughout the implementation period, with Pulse survey 3 receiving the most responses. This may reflect timing (teachers may have had more time to engage with the survey as it was later on in the term), length of completion window (it was open for longer than Pulse surveys 1 and 2) and the fact that we introduced a prize draw for a £50 voucher for Pulse survey 3, which may have incentivised completion.

Table 7: details on 'pulse' surveys sent to teachers during the trial

Pulse	Topics	Response rate
Pulse survey 1	<ul style="list-style-type: none"> <li>• Frequency of pre-teaching</li> <li>• Timing in relation to main lesson</li> <li>• Pupil engagement</li> </ul>	59/198 responses (24.8%)

<sup>11</sup> This was fewer than originally planned due to low response rates and clashes with the baseline and endline surveys, which were both open for longer than anticipated due to schools starting and finishing implementation at different points (some schools started and finished later).

	<ul style="list-style-type: none"> <li>● Ease of implementation</li> </ul>	
Pulse survey 2	<ul style="list-style-type: none"> <li>● Frequency of pre-teaching</li> <li>● Length of session</li> <li>● Timing in relation to main lesson</li> <li>● Challenges</li> <li>● Ease of implementation</li> </ul>	62/250 responses (26.1%)
Pulse survey 3	<ul style="list-style-type: none"> <li>● Frequency of pre-teaching</li> <li>● Timing in relation to main lesson</li> <li>● Additional support delivered to pupils in the comparison group</li> <li>● Pupil progress</li> <li>● Challenges</li> <li>● Ease of implementation</li> </ul>	117/238 responses (49.2%)

Overall, the pulse surveys (particularly pulse survey 3) provide useful insights into how pre-teaching was implemented and received in schools during the trial period, to supplement the retrospective data that was collected at endline. However, the information that was collected is limited by the following:

- The proportion of teachers who responded to all three surveys was low (18%) meaning there is the potential for bias in the results. In terms of future learning, this suggests that pulse surveys may be a less reliable way of collecting consistent monitoring data, in comparison with other methods.
- The pulse surveys were designed to act primarily as a ‘temperature check’ for teachers to report how they were finding the pre-teaching. As a result, they also had to be as short and simple as possible to encourage completion, which in turn limited the level of depth and detail of information collected. For example, we did not ask teachers to report how many sessions they completed and who attended each session, and so are unable to comment on how much pre-teaching pupils in the treatment group received (*dosage*).

#### Case studies

A case study approach was used to supplement findings from the teacher surveys, and to explore key areas of interest in greater depth. We conducted a total of six case studies during the delivery period (November 2024 to January 2025). In each case study we carried out an observation of a pre-teaching session, interviews with two members of staff, and a focus group with pupils allocated to receive pre-teaching (more information below).

A purposive sampling approach was used to recruit schools based on the following characteristics:

- **Implementation stage.** To reflect the fact that schools began delivery at different points in the term, and to ensure we captured data from both ‘on time’ and ‘late’ implementers.<sup>12</sup>
- **Proportion of pupils in the school eligible for FSM.** This is an indicator of socio-economic disadvantage, which has been identified as a potential moderating factor for the delivery and effectiveness of pre-teaching.

The achieved sample and breakdown of activities is shown in Table 8 below. We originally planned to use data on adherence and responsiveness from the pulse surveys to produce an implementation ‘score’ for sampling. This was not possible due to the limited response rate from the pulse surveys, and the fact that the pulse surveys were sent out later than planned. However, we managed to achieve a range of schools in terms of % of FSM pupils and whether they were ‘on time’ or ‘late’ in terms of implementation schedule.

Table 8: Case study sampling criteria

<sup>12</sup> This information was recorded during the trial, and used for sampling. It was hypothesised that this may reflect school engagement with the trial, and therefore, influence implementation (*fidelity/quality*).

School	FSM%	Implementation stage	Achieved activities
School 1	18.98%	On time	<ul style="list-style-type: none"> <li>● Observation of Year 1 pre-teach session</li> <li>● 1 x interview with project lead, 1 x interview with Year 1 session lead (teaching assistant)</li> <li>● Focus group with Year 1 pupils (n=6)</li> </ul>
School 2	27.10%	On time	<ul style="list-style-type: none"> <li>● Observation of Year 2 pre-teach session</li> <li>● 1 x interview with project lead, 1 x interview with Year 2 session lead (teacher)</li> <li>● Focus group with Year 2 pupils (n=6)</li> </ul>
School 3	24.70%	Late	<ul style="list-style-type: none"> <li>● Observation of Year 2 pre-teach session</li> <li>● 1 x interview with project lead, 1 x interview with Year 2 session lead (teacher)</li> <li>● Focus group with Year 2 pupils (n=5)</li> </ul>
School 4	8.80%	On time	<ul style="list-style-type: none"> <li>● Observation of Year 2 pre-teach session</li> <li>● 1 x interview with project lead, 1 x interview with Year 2 session lead (teacher)</li> <li>● Focus group with Year 2 pupils (n=6)</li> </ul>
School 5	35.70%	On time	<ul style="list-style-type: none"> <li>● Observation of Year 1 pre-teach session</li> <li>● 1 x interview with project lead, 1 x interview with Year 1 session lead (teacher)</li> <li>● Focus group with Year 1 pupils (n=6)</li> </ul>
School 6	66.10%	Late	<ul style="list-style-type: none"> <li>● Observation of Year 1 pre-teach session</li> <li>● 1 x interview with project lead, 1 x interview with Year 1 session lead (teacher)</li> <li>● Focus group with Year 1 pupils (n=4)</li> </ul>

### Observation of pre-teaching sessions

We observed one pre-teaching session per case study school. The observations were conducted using a semi-structured observation guide/checklist to capture adherence to the core components and indicators of quality. The guide/checklist was developed by the research team and was based on the core components of pre-teaching defined in the study plan (Bellier *et al.*, 2024). The document was designed as a ‘checklist’ where observers had to indicate whether each of the components had been delivered as intended and record details of any adaptations. The guide/checklist also included a section on ‘quality of delivery’. Indicators included delivery style, how the session was structured and organised, use of activities and resources, and the extent to which the session was tailored to pupil needs. For each indicator, observers recorded notes to capture additional details about the ways in which the session is delivered. Researchers supplemented this with more detailed field notes about the session structure, content (specific activities and techniques and how these were used), adaptations, the teachers’ delivery, and pupil engagement.

### Interviews with teachers

In each school we conducted two interviews: one with the school lead responsible for implementing the trial; and one with a member of staff (teacher or teaching assistant) who has delivered pre-teaching (a total of 12 interviews across the six case studies). The interviews lasted approximately 45 minutes each and were conducted remotely. The interviews were moderated with the use of a topic guide to ensure systematic coverage of key themes that address the evaluation research questions. Interviews were audio recorded.

### Focus groups with pupils

The aim of the focus groups was to understand pupils’ experience of the pre-teaching sessions, how the sessions had affected them (perceived outcomes), and what contributed to this (moderators and mediators). We conducted one focus group per school, made up of four to six pupils who were allocated to receive the pre-teaching sessions. These lasted between 10 and 20 minutes. The groups were segmented by year group to allow us to compare their experiences, and for

logistical reasons. In total, we conducted three focus groups with Year 1 pupils and three focus groups with Year 2 pupils. Teachers/teaching assistants were present, given the age of pupils, to support pupils to participate and maintain focus and order during the session. The presence of teachers/teaching assistants may have affected pupils' ability to speak honestly about the pre-teaching, especially as these staff members often delivered the sessions. However, this was outweighed by the potential challenges of conducting research with children of this age without a familiar adult. To mitigate this, researchers led the questioning, and teachers were briefed to intervene only for support with managing behaviour or distress.

A topic guide was designed to guide the discussion, and incorporated different activities to engage participants beyond simple verbal prompts/probes. This included the use of emotion cards, whereby participants could pick a card that best described how they felt in response to a prompt. The topic guide was used flexibly and adaptations were made based on learning. For example, we decided to keep the focus of our questions on the session that we had just observed, as this was easier for pupils to recall and discuss. Nevertheless, we encountered challenges with concentration and understanding, which were particularly prevalent in the focus groups with Year 1 pupils, and which are common when conducting such activities with pupils of this age. Therefore, while the focus groups provided some broad insights into how pupils were finding the pre-teaching and their level of engagement, there were challenges with eliciting responses, which limits the quality of the data, and means findings from the focus groups feature less prominently in the findings than other methods.

## Analysis

The qualitative and quantitative data collected through the IPE were analysed separately by different analysts to ensure both streams of data were thoroughly examined prior to comparing findings across sources. The aim of this was to prevent the results of one dataset from influencing the analysis of the other.

The teacher surveys were analysed using a combination of simple descriptive statistics, presenting the proportion of respondents choosing specific response options, and basic thematic analysis of open-text answers.

Data from the interviews, focus groups, and observations was managed and analysed by the same researchers who conducted them, using the Framework approach, which allows in-depth exploration of the data by case and by theme (Ritchie and Spencer 1994).<sup>13</sup> This consisted of creating a matrix in which to organise the data, based on the topic guides and observation proformas. Data was summarised and displayed in the matrix. This was followed by working through the managed data to draw out the range of behaviours, experiences, and views, while identifying similarities, differences, and links between them. Our analysis focused on identifying the range and diversity of experiences from across the entire sample (e.g. comparing and contrasting the views of teachers from across the six schools) as well as triangulating the different data sources from each case study to understand implementation in different contexts. The findings were then presented thematically rather than as narrative case studies.

Table 9: IPE methods overview

IPE dimension	Research questions addressed	Research methods	Data collection methods	Participants/data sources
Differentiation	2	Baseline survey	Online questionnaire	Trial lead and maths class teacher delivering pre-teaching
Differentiation, responsiveness, moderators, perceived impact	2, 4, 5, 7	Endline survey	Online questionnaire	
Fidelity, dosage, adherence,	1, 2, 3	Pulse surveys	Online questionnaire	Maths class teacher delivering pre-teaching
Fidelity, quality, adaptations, responsiveness,	1, 3, 4, 5, 6	Case studies (six schools)	Teacher interviews (total of 12, two per school)	Purposive sample of schools
			Pupil focus group (total of six, one per school)	

<sup>13</sup> An analytical framework was created with columns to represent the key subthemes and the rows representing individual participants interviewed or activities observed. Data was summarised in the appropriate cell.

perceived impacts, mediators, moderators,			Pre-teaching observation (total of six, one per school)	
---	--	--	---	--

## Timeline

Table 10: Timeline

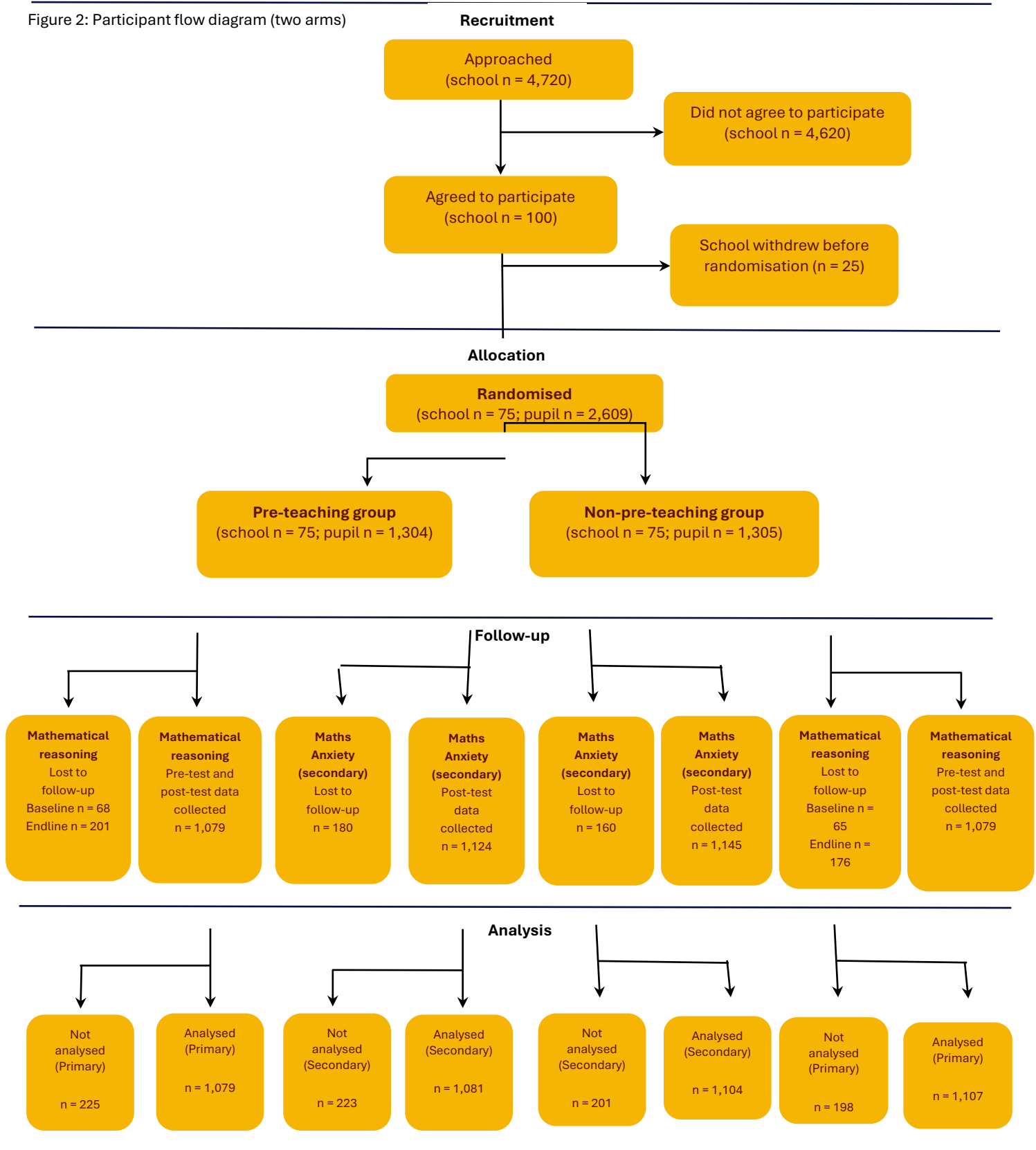
Dates	Activity	Staff responsible / leading
September 2022 – October 2022	Project set-up (scoping phase only)	BIT
October 2022 – January 2023	Initial scoping phase, including: <ul style="list-style-type: none"> <li>● Rapid evidence review</li> <li>● Interviews with teaching staff</li> <li>● Teacher Tapp surveys</li> </ul>	BIT
February 2023 – August 2023	Further in-depth scoping phase, including: <ul style="list-style-type: none"> <li>● Desk review to refine power calculations</li> <li>● Workshop with teaching staff</li> <li>● Teacher Tapp survey</li> <li>● Expert consultations</li> <li>● Trial design</li> </ul>	BIT
September 2023 – November 2023	Evaluation set-up <ul style="list-style-type: none"> <li>● Development and agreement of proposal for Teacher Guide pilot and full trials</li> <li>● Logic model development</li> <li>● First draft Teacher Guide</li> </ul>	The EEF/BIT
November 2023 – December 2023	Teacher Guide pilot set-up , including: <ul style="list-style-type: none"> <li>● Recruitment of schools for the pilot</li> <li>● Development of implementation materials</li> <li>● Development of evaluation materials</li> </ul>	BIT
January 2024	Ethical Approval for the trial (Teacher Guide pilot and full trial)	Senior Researcher at BIT
January 2024 – February 2024	Teacher Guide pilot and evaluation, including: <ul style="list-style-type: none"> <li>● Six schools implementing pre-teaching and providing feedback on pre-teaching and Guide</li> <li>● Qualitative evaluation through surveys, debrief workshop, and teacher diary notes analysis</li> </ul>	BIT
February 2024	Teacher Guide pilot reporting, including: <ul style="list-style-type: none"> <li>● Presentation of findings to the EEF</li> <li>● Development of final Teacher Guide version for the trial</li> </ul>	BIT
January 2024	Finalise recruitment documents for full trial, including: <ul style="list-style-type: none"> <li>● BIT website and the EEF webpage</li> <li>● Schools information sheet</li> <li>● MoU</li> <li>● Slides for calls with schools</li> </ul>	The EEF/BIT
January 2024 – August 2024	Study plan design	The EEF/BIT
February 2024 – July 2024	School recruitment to main trial	BIT
July 2024 – August 2024	Design baseline survey, pulse surveys, and observation guide/fidelity checklist	BIT
January 2024 – September 2024	Select and subcontract external assessors for test development and marking, including: <ul style="list-style-type: none"> <li>● liaising with Oxford Quantitative Reasoning team to develop the tests and training materials for teachers</li> <li>● liaising with Qa Research for secondary outcome test marking (CMAS-UK)</li> </ul>	BIT

Dates	Activity	Staff responsible / leading
August 2024	Selecting and subcontracting external partner for printing and distributing assessments	BIT
August 2024 – September 2024	Print and distribute baseline assessments	Secure and Confidential Documents (SCD)
September 2024 – October 2024	Administration of baseline survey and baseline testing (QRT only)	BIT
September 2024 – January 2025	Support schools to implement pre-teaching	BIT
September 2024 – November 2024	Schools administer and return baseline assessments	BIT/SCD
September 2024 – November 2024	Administer pulse surveys to teachers (September, October, November)	BIT
November 2024 – December 2024	Print and distribute endline assessments	SCD
December 2024 – March 2025	Schools administer and return endline assessments	BIT/SCD
January 2025 – February 2025	Conducts qualitative IPE activities, including: <ul style="list-style-type: none"> <li>• case study visits</li> <li>• interviews</li> <li>• surveys</li> </ul>	BIT
February 2025 – July 2025	Mark primary outcome (QRT) data, and return to BIT	Oxford
February 2025 – March 2025	Mark secondary outcome (CMAS-UK) data, and return to BIT	Qa Research
February 2025 – April 2025	IPE qualitative data analysis	BIT
February 2025 – June 2025	IPE quantitative data analysis	BIT
May 2025 – Aug 2025	Impact evaluation analysis	BIT
May 2025 – September 2025	Main reporting	BIT
September 2025	First draft report delivered to the EEF	BIT
October 2025 – November 2025	The EEF and peer review feedback on report	The EEF
March 2026	Report finalised and re-shared with the EEF	The EEF/BIT

## Impact evaluation results

### Participant flow including losses and exclusions

Figure 2: Participant flow diagram (two arms)



Approximately 4,720 schools were approached during the recruitment phase of the evaluation. Email was the most common method of contact, often using BIT’s network of schools, academy trusts, and local authorities from previous similar evaluations. Other means of communication were used also, such as paid advertisements within the Teacher Tapp smartphone application, dissemination via local authority contacts, and promotion on X (formerly Twitter) and LinkedIn by both the EEF and BIT. Of the 4,620 contacted schools that did *not* agree to participate, it is worth noting that the vast majority of schools simply did not reply to our outreach—the schools that did engage were on the whole receptive to One Step Ahead.

The study plan (Bellier *et al.*, 2024) power calculations were based on the conservative assumption that we would recruit only single-form entry schools (i.e. one class per year per school), implying a total (and maximum) of 24 pupils per school in the trial. However, we recruited and randomised 31 multi-form schools, which resulted in an average of 34.8 pupils per school, across 75 schools randomised. The minimum number of pupils—across both Year 1 and Year 2—is 21 pupils in one school, with a maximum of 78 pupils from one school.

As such, although this trial had relatively high levels of attrition, the study is still well-powered at the pupil level to detect a meaningful effect.

Table 11: MDES at different stages

		Design	Randomisation		Analysis	
		Overall	Overall	FSM	Overall	FSM
MDES		0.125	0.125	0.21	0.12	0.23
Pre-/post-test correlations	Level 1 (pupil)	0.29	0.29	0.29	0.54	0.46
	Level 2 (class)	–	–	–	–	–
	Level 3 (school)	–	–	–	–	–
ICCs	Level 2 (class)	N/A	N/A	N/A	N/A	N/A
	Level 3 (school)	N/A	N/A	N/A	N/A	N/A
Alpha		0.05	0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		Two-sided	Two-sided	Two-sided	Two-sided	Two-sided
Average cluster size		N/A	N/A	N/A	N/A	N/A
No. of schools	Total <sup>a</sup>	95	75	75	74	74
No. of pupils	Pre-teaching Group	1,140	1,303	369	1,107	304
	Comparison group	1,140	1,305	363	1,079	306
	Total:	2,280	2,608	732 <sup>b</sup>	2,186	610

<sup>a</sup> The surveys were kept open for longer than planned to give schools that started implementing pre-teaching later an opportunity to complete them. Note that this trial was a within-school randomisation at the pupil level.

<sup>b</sup> Please note we are missing some data on FSM status (n=171).

N/A=not applicable.

## Attrition

Table 12: Pupil-level attrition from the trial (primary outcome)

		Pre-teaching group	Comparison group	Total
No. of pupils	Randomised	1,304	1,305	2,609
	Analysed	1,079	1,107	2,186
Pupil attrition (from randomisation to analysis)	Number	225	198	423
	Percentage	17.3	15.2	16.2

Attrition in the trial (n = 423, 16.2%) arose from two main sources: pupil withdrawals (n = 262, 10.0%) and missing test results (n = 161, 6.2%). Withdrawals include whole-school, class/year level (i.e. one class, or all classes within a year, without the whole-school withdrawing), and individual-level withdrawals. Four schools formally exited the trial at various points. The reasons offered for this were broadly a combination of staff or leadership changes and other unanticipated changes that reduced teaching staff's ability to take part in the initiative, and/or that the teacher-administered assessments (for Year 1) were found to be too time-consuming and difficult to administer. This whole-school attrition accounts for 128 of 262 pupil withdrawals, while the remaining 134 withdrawals were recorded at the pupil level. Of the remaining 134 withdrawals, 48 were entire classes or year groups (all Year 1) rather than specific individuals withdrawing, and nine other pupils withdrew because they left their school during the trial.<sup>14</sup> Schools collectively did not provide us with reasons behind the remaining 77 pupil-level (i.e. individual pupil) withdrawals.

The 161 non-withdrawn pupils excluded from the analysis were dropped due to incomplete QRT data: 51 had missing baseline scores and 115 had missing endline scores, with only five pupils missing both outcomes, thereby compounding attrition. 43% of baseline missingness was due to pupils being absent on the day of the test while the remaining cases were unexplained or reflected a pupil's choice not to sit the test.

## Pupil and school characteristics

Table 13 below summarises the baseline pupil-level characteristics of pupils in the pre-teaching and comparison group as randomised. In general, it shows a very high level of balance between groups, with negligible differences. It also shows that, compared to the national pupil population in the same age group, the trial had a substantially higher proportion of pupils eligible for FSM (30% vs 25.7%). It also had a higher proportion of pupils with EAL (31.5% vs 23.4%).

Overall, the sample appears well-balanced across pupil characteristics. Normalised differences with a magnitude of 0.1 or less indicate a negligible correlation between the covariate and assignment to treatment group, which can usually be addressed through covariate adjustment in the regression (Austin 2009, p. 1233). Following this interpretation of the magnitude of differences, both the randomised and analysed sample appear to be well-balanced on FSM status, gender, EAL status, and year group.

We further stress-check the difference in the pre-test score (QRT baseline) between pupils in the pre-teaching group versus the comparison group and find that the difference between these baseline scores are also negligible with a standardised difference of 0.07 (95% CI: -0.01, 0.15).

Overall, we do not expect these minor differences between the two groups to introduce any bias into the primary analysis but, as pre-specified, we re-run the primary, secondary, and subgroup controlling for FSM status, EAL status, gender, and year group in an additional analysis.

<sup>14</sup> Class withdrawals are typically instances where a particular teacher is facing challenging conditions (e.g. new member of staff who was not present when the school signed up for the project the previous year, teacher who no longer has teaching assistant support), and the project lead requested that the class be withdrawn. We usually accommodated those requests, balancing it out against the risk of the whole-school withdrawing if we did not.

Table 13: Baseline characteristics of groups as randomised

School level (categorical)	National level mean	Pre-teaching group		Comparison (No pre-teaching) group	
		n/N (missing)	Count (%)	n/N (missing)	Count (%)
Type of school: Multi-form		31 (0)	41.3	31 (0)	41.3
Type of school: Single-form		44 (0)	58.7	44 (0)	58.7
School level (continuous)		n/N (missing)	Mean (SD) <sup>a</sup>	n/N (missing)	Mean (SD)
Percentage FSM	25.7 <sup>b</sup>	74 (1)	31.1 (23.5)	73 (2)	29.9 (22.5)
Percentage EAL	23.4	74 (1)	31.0 (29.5)	74 (1)	29.5 (28.8)
Pupil level (categorical)		n/N (missing)	Count (%)	n/N (missing)	Count (%)
Gender: Male	51.2% <sup>c</sup>	606 (59)	48.7	572 (58)	45.9
Gender: Female	48.8%	639 (59)	51.3	675 (58)	54.1
FSM: Yes	25.7% <sup>d</sup>	369 (86)	30.3	363 (85)	29.8
FSM: No	74.3%	849 (86)	69.7	857 (85)	70.2
EAL: Yes	23.4% <sup>e</sup>	400 (56)	32.1	387 (52)	30.9
EAL: No	76.6%	848 (56)	68.0	866 (52)	69.1
Year 1	49.3% <sup>f</sup>	633 (0)	48.5	629 (0)	48.2
Year 2	50.7%	671 (0)	51.5	676 (0)	51.8
Pupil level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)
Baseline		1,236 (68)	-0.03 (0.98)	1,240 (65)	0.03 (1.02)
Endline		1,103 (201)	-0.01 (1.02)	1,129 (176)	0.01 (0.98)

<sup>a</sup> These school-level figures indicate the average and SD measures within each treatment group, at the school level. For example, in the pre-teaching group there are, on average, 31.1% pupils in the school who qualify for EAL status with an SD of 23.5. This is in contrast to, on average, 29.9% of pupils in the school who qualify for EAL status with an SD of 22.5. Randomisation was done within-school and as such these figures should not be interpreted as a balance check across samples and instead, only as descriptives.

<sup>b</sup> Across all pupils in England, 25.7% were eligible for FSM in the academic year of 2024/2025. This data can be further broken down into 21.8% of Year 1 pupils and 24.5% of Year 2 pupils. (DfE, 2025b).

<sup>c</sup> For academic year 2024/2025, Years 1 and 2 only (GOV.UK, 2025).

<sup>d</sup> For academic year 2024/2025 (DfE, 2025a; DfE, 2025b).

<sup>e</sup> Ibid.

<sup>f</sup> For academic year 2024/2025, Years 1 and 2 only (GOV.UK, 2025).

Table 14: Balance checks of selected variables at randomisation

	Pre-teaching group		Comparison group		Normalised difference
	Mean	SD	Mean	SD	
FSM	0.30	0.46	0.30	0.46	0.01
Male	0.49	0.50	0.46	0.50	0.06
EAL	0.32	0.47	0.31	0.46	0.03
Year 1	0.49	0.50	0.48	0.50	0.01
Baseline QRT score	-0.03	0.98	0.03	1.02	0.06

Table 15: Balance checks of selected variables (primary analysis)

	Pre-teaching group		Comparison group		Normalised difference
	Mean	SD	Mean	SD	
FSM	0.29	0.46	0.29	0.45	0.01
Male	0.48	0.50	0.46	0.50	0.04
EAL	0.33	0.47	0.30	0.46	0.05
Year 1	0.47	0.50	0.47	0.50	0.01
Baseline QRT score	-0.02	0.99	0.05	1.03	0.07

## Outcomes and analysis

The impact evaluation addresses three research questions, as listed below. Research question 1 is the primary research question and research questions 2 and 3 are secondary research questions. Research questions 1 and 2 are reported in the ‘Primary and secondary analysis’ section below. Research question 3 is reported in the [‘Subgroup analyses’ section](#).

1. Was there a difference in pupils’ mathematical reasoning skills at the end of the Autumn Term (as measured by the QRT) between pupils who were allocated to receive the pre-teaching, and pupils in the comparison group?
2. Was there a significant difference in pupils’ anxiety related to maths at the end of the Autumn Term (as measured by their score on the CMAS), between pupils who were allocated to receive the pre-teaching and pupils in the comparison group?
3. Are there heterogeneous effects by FSM status? By baseline attainment (for pupils’ mathematical reasoning skills only)?

Table 16 presents the means and SDs of the primary and secondary outcome measures, including all pupils who did not withdraw from the trial and for whom we have endline data.

Table 16: Raw means and SDs of outcome measures

Outcome	Pre-teaching group		Comparison group	
	n	Mean (SD)	n	Mean (SD)
QRT endline score (primary outcome) <sup>a</sup>	1,103	-0.04 (0.99)	1,129	0.03 (1.02)
CMAS-UK (secondary outcome) <sup>b</sup>	1,078	1.90 (0.35)	1,103	1.89 (0.35)

<sup>a</sup> The primary outcome analyses controls for the QRT score at baseline.

<sup>b</sup> The CMAS-UK score has a theoretical minimum of 1 and a maximum of 3.

### Primary and secondary analysis

As outlined in the [‘Methods’ section](#), we run an OLS regression with school-level fixed effects on the sample of all pupils randomised. We exclude pupils who have withdrawn from the trial and pupils for whom we do not have baseline and/or endline data. The results from the OLS regression are in [Appendix G](#), and a summary of both analyses are provided in Table 17 below.

Table 17: Primary and secondary analysis

Outcome	Unadjusted means				Effect size		
	Pre-teaching group		Comparison group		Total n (Condition 1 and 2)	Hedges’ g (95% CI)	P-value
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
QRT endline score	1,079 (225)	-0.01 (-0.07, 0.05)	1,107 (198)	0.01 (-0.05, 0.07)	2,186 (423)	0.01 (-0.07, 0.09)	0.80
CMAS-UK	1,078 (223)	1.89 (1.87, 1.91)	1,103 (201)	1.89 (1.87, 1.91)	2,181 (424)	0.01 (-0.07, 0.10)	0.76

Table 17 presents the results of the primary and secondary analyses. To ensure consistency and comparability of results, the QRT scores (outcome) were standardised for analysis. This is a deviation from the study plan (Bellier *et al.*, 2024), however, it was deemed necessary as the QRT tests administered to Year 1 and Year 2 pupils had different score ranges (e.g. Year 1 baseline out of 23 and endline out of 24, while Year 2 was out of 21 for both baseline and endline). The endline QRT test for each year group was also designed to be more challenging than its baseline counterpart. Please see the [‘Methods’](#) section for further detail. As a result, all group means and regression estimates are expressed in SD units relative to the pooled sample mean.

#### Primary analysis

The standardised, unadjusted mean for the primary outcome (QRT endline score) is -0.10 in the pre-teaching group and -0.10 in the comparison group. These mean scores are the treatment estimates from an unadjusted model with treatment assignment as the only covariate. The regression coefficient from the primary analysis including the baseline score as a covariate, suggests a very small (negligible) positive effect (beta = 0.009) of the treatment (pre-teaching) on pupils’ mathematical reasoning skills. While this is the best estimate of pre-teaching’s impact, the 95% CIs of the estimate suggest the true effect could range from a Hedges’ g of -0.07 (small negative effect) to 0.09 (small positive effect). The wide range of CIs imply some uncertainty around this result. Teachers were instructed to identify lower-attaining pupils for pre-teaching. However, we would expect QRT baseline scores to be similar across the two groups as pupils were randomised into treatment (pre-teaching) and comparison (no pre-teaching) groups. The baseline assessments were also conducted prior to randomisation by BIT. However, at baseline, the pre-teaching group had an average score of -0.04 (standardised) points (95% CI: -0.09–0.02; SD = 0.98), compared with 0.03 points (95% CI: -0.03–0.09; SD = 1.02) in the comparison group. This is unlikely to be due to differences in test administration, given that teachers were blind to treatment group allocation at the time when they delivered the baseline test (randomisation was done after baseline testing). By the endline assessment, both groups reached similar average scores (pre-teaching: -0.01, 95% CI: -0.07–0.05; comparison: 0.01 95% CI: -0.05–0.07), but the pre-teaching cohort exhibited a somewhat larger gain (mean increase of 0.03 points vs 0.02 points), suggesting that pre-teaching may have had a positive impact on pupil progress. However, as mentioned above, the estimated effects are also compatible with a null or negative effect of pre-teaching on mathematical reasoning as measured by the QRT.

#### Secondary analysis

For the secondary outcome, pupil mathematical anxiety, measured by the CMAS-UK average score, the regression analysis suggests a negative effect (beta = 0.004) of the treatment (pre-teaching) on pupils’ mathematical anxiety scores, namely, pupils in the treatment group were *more* anxious than those in the comparison group. While this is the best estimate of pre-teaching’s impact, the 95% CIs suggest the true effect, expressed in Hedges’ g, could range from -0.07 (slightly lowered anxiety in the treatment group), to 0.10 (slightly higher anxiety in the treatment group). As with the primary outcomes, the

range of CIs imply that there is significant uncertainty around this result, and that the findings are also compatible with a null or positive effect of pre-teaching on pupils' anxiety related to maths. As this regression did not include any covariates, we derive the unadjusted mean CMAS-UK scores from the same regression, which indicates a score of 2.11 for the treatment group and 2.10 for the comparison group.

In summary, neither the primary nor the secondary outcome measures show a conclusive positive or negative impact of the pre-teaching, based on these regression results. Our best estimate is that the pre-teaching had no impact on pupils' mathematical reasoning skills, nor their anxiety towards maths.

### **Analysis in the presence of non-compliance**

Compliance and overall fidelity data were collected in two surveys: i) the endline survey, administered at the end of the delivery of pre-teaching; and ii) the final pulse survey, administered at November 2024. In total, 60 out of 75 schools randomised (80%) provide at least one teacher's response to the endline survey. Further details on this survey are reported in the ['IPE results' section](#).

For this trial, we established the minimum level of compliance at the teacher level, using three criteria as measured by the endline survey based on teachers' retrospective assessments. Specifically, a teacher was required to report compliance with all three of the following criteria:

- **Criteria 1.** Teachers (or their teaching assistant) have delivered maths pre-teaching sessions at least twice a week during the Autumn Term (unless unforeseen circumstances occur).
- **Criteria 2.** Sessions were delivered to the group of pupils allocated to receive the sessions.
- **Criteria 3.** This group stayed fixed during the Autumn Term.

Criteria 1 was measured by teachers' response to the question: 'How often have the pre-teaching sessions taken place, on average?' where teachers who select either 'Twice a week' or 'More than twice a week' are classified as compliant and teachers who select lower frequency are classified as non-compliant.

Criteria 2 and 3 were measured interrelatedly based on teachers' response to two questions. The endline survey first asks teachers: 'Have there been any changes in the group of pupils receiving the pre-teaching?'. If teachers select 'Yes, one or more pupils joined the group', they are shown the question: 'Were any of the pupils who joined the pre-teaching group initially assigned to the comparison (no pre-teaching) group, in the list we shared with you?' Teachers who indicate that a pupil from the comparison group joined their pre-teaching sessions are assigned a non-compliant value for *both* Criteria 2 and 3. Conversely, teachers who indicate that groups remained consistent or that a pupil left the group, are classified as compliant with both Criteria 2 and 3.

We then categorised all responses as 'compliant' (all Criteria 1, 2, and 3 met), responses meeting two of the three criteria as 'partially compliant', and responses meeting fewer than two criteria as 'non-compliant'. In Table 18, we report the proportion of fully, partially, and non-compliant teachers, and a breakdown of compliance to each of the three criteria listed above.

A large proportion of respondents report high compliance with pre-teaching delivery: the vast majority fully met all three criteria (78%) and a further 21% met two out of three criteria. Only a very small minority were compliant with just one criteria (2%), and we received no reports of non-compliance, namely, compliance with no criteria. This can be further broken down into the three individual compliance criteria: 81% stated that they delivered pre-teaching twice a week; 97% stated that it was delivered to allocated pupils; and 98% stated that the groups were fixed throughout the Autumn Term. Overall, compliance scores as per the endline survey are heavily skewed towards full compliance, with little variability across criteria. These findings are further discussed in light of the IPE findings below.

The endline survey sample had fairly good coverage across schools. We received at least one teacher response from 60 out of 71 schools (85%) and we received more than one teacher response from most schools in our sample (41 out of 60: 68%). We report compliance below at the teacher level there is some variation within schools. Out of the 60 schools that did respond, we find 36 schools report full compliance across all teacher responses.

In the study plan (Bellier *et al.*, 2024), we pre-specified that we would use teacher-level indicators of compliance to derive pupil-level compliance and use this to derive a CACE estimate. However, as outlined in the '[Methods](#)' section, we do not proceed with this analysis as we are unable to link teacher's responses to pupil's responses. This implies that we are not able to assess the effect of compliance on the reported effect on the primary outcome.

Table 18: Compliance results from the endline survey (response level)

Compliance	Number of responses (N = 126)	% of responses
Fully compliant (3 out of 3)	98	78
Partially compliant (2 out of 3)	26	21
Partially compliant (1 out of 3)	2	2
Non-compliant (0 out of 3)	0	0

This is broken down into compliance with each criterion in Table 19 below.

Table 19: Compliance results from the endline survey (response level)

Compliance	No. of responses (N = 126)	% of responses <sup>a</sup>
Pre-teaching twice a week (1)	102	81
Delivered to allocated pupils (2)	122	97
Group stayed fixed (3)	124	98

<sup>a</sup> Note that this is the % of respondents who indicate compliance with this criteria and therefore, rows are not mutually exclusive nor display column-wise percentages.

We further group responses by schools that respond to the endline survey (n= 60), and take an average compliance score as shown in Table 20 below.

Table 20: Compliance results from the endline survey (school level)

Compliance	No. of schools (N = 60)	% schools
Comply with all three criteria, on average	36	60
Comply with at least 2.5 criteria, <sup>a</sup> on average	50	83
Comply with at least two criteria, on average	60	100

<sup>a</sup> Please note that there were no 0.5 measures of compliance. The compliance levels are reported at the school level and 'Comply with at least 2.5' implies that across the school, the average compliance score (1, 2, or 3) reported by teachers is at least 2.5.

## Missing data analysis

Our final analysis sample is missing 133 covariate data observations (QRT baseline scores), making up 5% of the randomised sample. Similarly, the sample at analysis is missing 377 outcome data observations (QRT endline scores), constituting 14% of the data. There are 2,186 complete observations for the primary analysis.

As specified in the study plan (Bellier *et al.*, 2024), we provide a breakdown of the missingness across covariate and outcome for the primary outcome analysis between treatment and comparison groups in Table 21 below.

Table 21: Missing outcome and covariate data (by treatment group)

Measure	Pre-teaching group		Comparison group	
	n missing	Count (%)	n missing	Count (%)
Outcome (QRT endline score)	201	15.4	176	13.5
Covariate (QRT baseline score)	68	5.2	65	5.0

As the covariate data approaches the 5% threshold and outcome data crosses the threshold, we follow the study plan (Bellier *et al.*, 2024) specification to determine whether data are MAR or MNAR. MAR is when missingness in the data is correlated with other covariates while MNAR is when missingness is correlated with any unobservable characteristics.

### Missing covariate data

The pre-test score is the only covariate in the main analyses. In line with the study plan (Bellier *et al.*, 2024), we conduct an analysis to determine whether the MAR i.e., it is predicted by any of the observed covariates. To do this, we create a binary variable where the value 1 is assigned if any of the QRT baseline score is missing and the value 0 if it is present. We then run a logistic regression to investigate whether this binary indicator of completeness is predicted by any of the following covariates: gender; year group; EAL status; FSM status; and treatment allocation. The further appendices (e.g. [Appendix F](#)) contains the full regression output. The analysis suggests that the QRT baseline score is significantly more likely to be missing for pupils who have FSM status. Specifically, 3.3% non-FSM pupils are missing a QRT baseline score and 5.7% of FSM pupils are missing a QRT baseline score. This suggests that the baseline data are MAR. We do not believe that this warrants further analysis (e.g. multiple imputation) for two main reasons: i) the level of missingness is close to, but does not greatly exceed the 5% threshold; and ii) the final sample used for analysis remains well-balanced on FSM status, minimising the potential for bias (see Table 15).

### Missing outcome data

The study plan (Bellier *et al.*, 2024) specified that we will further investigate data missingness if this is predicted by observed [covariates](#). To test this we create a binary variable where the value 1 is assigned if the QRT endline score is missing and the value 0 if it is present. We then run a logistic regression to investigate whether this binary indicator of outcome completeness is predicted by any of the following covariates: gender; year group; EAL status; FSM status; and treatment allocation. The further appendices (e.g. [Appendix F](#)) contains the full regression output. The analysis suggests that outcome data is significantly less likely to be missing for pupils who are in Year 2. In the study plan (Bellier *et al.*, 2024), we specify that if any of the observed covariates predicts missingness, we will run a complete case analysis with and without covariates and compare the results. We do not specify any analyses for the scenario we face: covariates that are not included in the analyses predicting missingness. However, given that this covariate is not included in the main primary and secondary analysis however, it does not change the main results and we therefore, do not proceed with this additional sensitivity check. We also further note that we run two models to estimate the impact of pre-teaching on endline scores: i) the primary analysis, which only uses the QRT baseline score as a covariate; and ii) an additional analysis with all covariates (including year group). The estimates of impact do not change with the addition of these covariates.

## Subgroup analyses

Table 22: Primary analysis for FSM subgroup (two methods)

Outcome	Unadjusted means				Effect size		
	Pre-teaching group		Comparison group		Total n (Condition 1 and 2)	Hedges' g (95% CI)	P-value
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
QRT endline score (Interaction model)	1,041 (263)	-0.03 (-0.10, 0.03)	1,067 (23)	-0.00 (-0.06, 0.06)	2,108	-0.10 (-0.25, 0.06)	0.22
QRT endline score (FSM-only model)	304 (0)	-0.17 (-0.28, -0.06)	306 (0)	-0.12 (-0.23, -0.02)	610	-0.09 (-0.25, 0.07)	0.28

The study plan (Bellier *et al.*, 2024) specifies a subgroup analysis to explore whether the impact of the pre-teaching varied based on pupils' FSM status (Equation 4). We used two approaches for this: one model run on the subsample of FSM-eligible pupils; and another on the full sample including an interaction term for FSM eligibility. The dependent variable for both analyses was the QRT endline score.

As Table 22 indicates, when estimating the effect on the FSM subgroup only by restricting the sample, those who were eligible for FSM suggest a very small (negligible) negative effect (beta = -0.07). While the best estimate of pre-teaching's impact on FSM pupils is a small negative effect (Hedges' g = -0.09) the true effect may range between a moderate negative effect and a small positive effect (95% CI of Hedges' g: -0.25, 0.07).

In a second model, we included an interaction term. The coefficient on the interaction term (interpreted as the difference in effects between FSM and non-FSM pupils) is small (see Hedges' g for the interaction effect model). Our best estimate of the difference in effects between the two groups is a small negative effect, however, the CI on this estimate is fairly wide (ranging from a moderate negative effect to a small positive effect).

The analysis conducted using an interaction term, and the analysis specifically on the subsample of FSM-eligible pupils both suggest that the *additional* effect on FSM pupils is slightly negative. While our best estimate of this additional impact on FSM pupils, the CIs for this estimate are fairly wide and suggest that this result should be interpreted carefully as the true additional effect on FSM pupils may range between a moderate negative effect to a small positive effect. Overall, the analysis does not provide sufficient evidence to make conclusive statements about the impact of pre-teaching on FSM pupils, and we caution against the interpretation of a negative effect.

## Additional analyses and robustness checks

Table 23: Primary analysis for FSM subgroup (two methods)

Outcome	Unadjusted means				Effect size		
	Pre-teaching group		Comparison group		Total n (Condition 1 and 2)	Hedges' g (95% CI)	P-value
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
QRT endline score (Primary outcome with covariates)	1,061 (243)	-0.01 (-0.07, 0.05)	1,084 (221)	0.01 (-0.05, 0.06)	2,145	0.01 (-0.07, 0.10)	0.77
CMAS-UK (Secondary outcome with covariates)	1,059 (245)	1.89 (1.87, 1.91)	1,080 (225)	1.89 (1.87, 1.91)	2,139	0.004 (-0.08, 0.09)	0.94
CMAS-UK (Secondary outcome with QRT score as covariate)	1,053 (225)	1.89 (1.87, 1.91)	1,082 (225)	1.89 (1.87, 1.91)	2,135	0.01 (-0.07, 0.10)	0.73
QRT endline score (FSM interacted with treatment, with covariates)	1,030 (274)	-0.03 (-0.10, 0.03)	1,053 (252)	-0.01 (-0.07, 0.05)	2,083	-0.10 (-0.25, 0.06)	0.22
QRT endline score (‘On time’ indicator interacted with treatment)	1,079 (225)	-0.01 (-0.07, 0.05)	1,107 (198)	0.01 (-0.05, 0.07)	2,186	0.10 (-0.03, 0.23)	0.13

As specified in the study plan (Equation 5) (Bellier *et al.*, 2024), we perform a subgroup analysis to study potential heterogeneous effects based on whether schools started implementing the pre-teaching ‘on time’ or ‘late’, reflecting unanticipated variation in the timing of implementation during the trial. The regression analysis suggests that pupils who were in schools that delivered the pre-teaching ‘on time’ scored slightly higher than pupils in the ‘late’ group, estimated to be approximately 0.10 (95% CI: -0.03, 0.23) SDs higher (Table 23).

The QRT scores, broken down by Year 1 and Year 2, suggest that patterns of scores may have differed between ‘on time’ and ‘late’ schools, particularly for Year 1 pupils. Among Year 1 pupils, those in schools that administered assessments on time scored higher than those in schools that administered them late—by 0.47 SD higher in the pre-teaching group and 0.22 SD higher in the comparison group. This indicates that while timing effects were evident for all pupils, the difference was larger for pre-teaching pupils, suggesting that any treatment effect may have been stronger where assessments were delivered on time. However, for Year 2 pupils, the differences between ‘on time’ and ‘late’ schools were minimal (0.01 SD and 0.00 SD respectively). Taken together, these findings suggest that the delivery timing may have moderated the impact of pre-teaching—mostly in Year 1—even though the formal regression results were not conclusive. This is explored further in relation to the IPE findings in the ‘Conclusion’ section below.

As a robustness check, we also re-ran the primary and secondary analyses with additional covariates including year group, gender, and EAL status. The primary analysis with additional covariates echo the main analysis (only baseline score as a covariate) exhibits a similar effect size, suggesting that the effect of pre-teaching does not change from the primary analysis after accounting for these demographic characteristics. The secondary outcome analyses similarly indicate a negligible effect of pre-teaching on pupils’ mathematical anxiety mirroring the main analyses.

## Implementation and process evaluation results

This section describes how pre-teaching was implemented in the trial (the research questions we set out to address in the IPE can be found in the [‘Evaluation objectives’](#) section). It covers adherence to the core components of the pre-teaching, quality, pupil engagement; and usual practice/differentiation. This is followed by a discussion of the perceived outcomes of the approach and key mechanisms/moderators in relation to pre-teaching’s logic model.

### Adherence

Overall, the results of the quantitative [compliance analysis](#)<sup>15</sup> suggest that teachers who responded to the survey adhered closely to the core components set out in the logic model and Teacher Guide.<sup>16</sup> We present results for each component in turn, describing the factors that influenced adherence and teachers’ experience of implementation.

### Frequency

The survey results suggest that most teachers were able to deliver regular pre-teaching in line with the guidelines. At endline, 81% of respondents indicated that pre-teaching sessions had been delivered at least twice a week during the Autumn Term. This was similar in Pulse survey 3 (76%) but lower in Pulse surveys 1 and 2 (51% and 65%, respectively). While this may reflect some inconsistency in frequency of pre-teaching over the trial period (perhaps indicating a ‘warm-up phase’), these results should be interpreted with caution as the sample size was low and likely to be subject to selection bias (see [‘Methods’](#) section).

The following factors, which were interlinked, are likely to have influenced how often teachers delivered the pre-teaching:

1. **Teacher buy-in.** The teachers/teaching assistants we interviewed were enthusiastic about the trial and motivated to deliver pre-teaching. They expressed an interest in its potential to support lower-achieving pupils and boost confidence. Consequently, it was something that some teachers/schools had already considered or were planning to introduce.
2. **Ease of implementation.** As a Teacher Choice trial, pre-teaching was intended to be relatively easy to adopt without extensive preparation, resources, or time commitment. Overall, the survey data suggests that this was true for many teachers—71% rated implementation as ‘easy’ in the endline survey. They attributed this to the experience of teaching staff (57%), having access to physical resources (57%), and the sessions being short in length (46%). The latter came through strongly in the interviews—teachers said if the sessions were longer, finding time to fit them in would have been difficult. This suggests that increasing the length of pre-teaching sessions, in an attempt to increase any positive impact on outcomes, is not a viable option for schools. Interviewees also noted that because it was not a prescriptive approach, it required minimal extra work beyond their usual lesson planning. This is supported by the survey, which suggests planning was not overly intensive (71% said they spent between 5 and 20 minutes planning each pre-teaching session, 20% said it took them less than five minutes.)
3. **Timetabling.** This was something that many teachers struggled with, despite the short length of the sessions. Nearly three-quarters (72%) of endline survey respondents said that it was challenging to find time in the day to deliver the sessions prior to the maths lesson. In the interviews, teachers described their timetables as ‘tightly packed’, with limited free time for additional activities. Initially, this led to delays with teachers starting the pre-teaching and committing to regular delivery, according to some project leads we interviewed. In the Pulse survey 1, 39% of respondents said they had not yet started delivery.

<sup>15</sup> To keep data collection requirements to a minimum, compliance was measured at the teacher level using responses to the endline survey. We did not ask teachers to record information on attendance and number/length of sessions. The survey response rate was 46%. We therefore, do not have compliance data for teachers in just over half of schools. There is a risk that the surveys may provide a biased picture of teacher compliance.

<sup>16</sup> These components were: frequency (twice weekly); duration/timing (approximately 15 minutes, prior to the main maths lesson); recipients (a fixed group of pupils assigned to the treatment group); and relationship to the main lesson (forward-looking focus).

Nevertheless, at endline, most teachers (81%) reported delivering the sessions at least twice a week, which suggests that teachers found ways around this.<sup>17</sup> A key strategy was to formally build pre-teaching into the timetable. Using time reserved for assemblies and collective worship worked well. However, even with this approach, challenges remained:

- Finding enough suitable time slots was difficult, especially for multi-form entry schools that needed to schedule two sessions per class each week to meet the requirement.
- Changes to timetables (e.g. due to whole-school events) caused some sessions to be cancelled. This may have reduced the total amount of pre-teaching that pupils received, especially given the short delivery period (one term).

It is also possible that the endline survey results overestimate how many teachers delivered the sessions twice a week. There are two possible reasons for this. There may have been a selection bias, with more compliant teachers responding to the survey (the sample was small, 46% response rate) or teachers' may have simply overestimated the frequency of delivery due to recall bias.

4. **Staffing.** This was another important practical consideration for schools, which affected their ability to deliver the sessions twice a week. Just over half of endline survey respondents (56%) said that staff capacity was a challenge. Interviewees, particularly class teachers, said they had to miss sessions when they lacked cover (e.g. due to staff absence) or because they had to handle other urgent priorities. Despite these difficulties, teachers valued the opportunity to deliver small group support, which they would rarely otherwise get.

Teaching assistants were involved in delivering pre-teaching sessions to mitigate teacher resource challenges. Both our recruitment materials and the Teacher Guide stated that while the sessions would ideally be delivered by the class teacher, it was possible for a teaching assistant to either help with, or directly deliver, the pre-teaching sessions where this was not possible. This built on findings from our scoping phase, and Teacher Guide pilot, which found that giving schools flexibility on this (i.e. allowing teachers to rely on teaching assistant support if and when needed) was seen as positive, provided that the Teacher Guide contained enough guidance (which led us to include more examples than beforehand). The endline survey showed there was considerable variation in the number of teaching assistants involved in delivering pre-teaching, which ranged from 0 to 9. The average number was 2.

While using teaching assistants helped to ensure the sessions got delivered, the use of teaching assistants in some schools might have moderated the impact of pre-teaching, and could help to explain the lack of treatment effect. In the interviews, some project leads felt that teaching assistant-led pre-teaching could be less effective than teacher-led sessions. This was primarily attributed to teachers' stronger subject knowledge and deeper understanding of the curriculum and learning objectives. Effective teaching assistant delivery was reliant on teachers creating clear plans and instructions for teaching assistants, and teaching assistants being able to deliver the content as intended and provide feedback on pupil understanding. There were examples of where this had been done well in the case studies, for example, where a teacher created a clear plan and set of instructions for the teaching assistant, discussed this with them and asked the teaching assistant to provide written feedback on how the session went. However, it is uncertain whether and how much this happened across the whole sample of participating schools. Furthermore, project leads also noted that less experienced teaching assistants were less confident in leading the sessions (discussed further below in 'Quality' section). The EEF's evidence-based guidance on the deployment of teaching assistants clearly lays out the challenges of using teaching assistants effectively. Specifically, the EEF recommends that teaching assistants are not used to replace teachers, and that pupils who struggle most should spend more time with their teacher (not their teaching assistant) as compared to their peers (EEF, 2025a). The EEF's guidance does not rule out using teaching assistants to deliver well-chosen, evidence-based, structured programmes where appropriate, but it may be that this particular type of approach is not appropriate for teaching assistants, especially in resource-constrained schools, where they may not receive the support needed to deliver it well (EEF, 2025a).

---

<sup>17</sup> It is worth noting however, that teachers may not have found this sustainable for a full school year, especially given concerns around staffing (see below).

## Duration and timing

The Teacher Guide advised running the sessions for around 15 minutes, but the exact length was left up to teachers. Around 73% of endline survey respondents said that the sessions lasted 10–20 minutes on average. However, a quarter (25%) reported that their sessions were shorter (less than ten minutes on average). In the interviews, teachers whose sessions were shorter said that it was difficult to keep pupils engaged for longer than this due to their limited attention spans. This was particularly the case for Year 1 pupils, as illustrated by the following quote:

*I think any longer would have felt like a lesson to them if I'm being honest. I thought, especially year one, getting them to sit on the carpet for any length of time [is a challenge]...any longer I think may have got in the way of what we wanted from them. (Session lead, Case study school)*

However, sessions could also be inadvertently shorter than planned due to the timetabling and staffing constraints discussed above. Teachers commented that sessions could be 'rushed' as a result. The short duration of some pre-teaching sessions may have prevented teachers from meaningfully introducing key concepts and addressing misconceptions. This limitation could have been a barrier to improving pupils' understanding of the main lesson and help to explain why pre-teaching had no overall effect.

The timing of sessions (proximity to the main lesson and time of day) was left up to schools. Around half of respondents (53%) said they delivered pre-teaching sessions on the same day as the main lesson. Around 46% said they happened earlier in the same week. Some teachers believed that same day pre-teaching was better for pupil recall but timetables did not always allow for this

*Finding the time to fit it in is difficult. Our maths lessons are first thing in the morning so any pre-teaching has to be done the day before. (Pulse survey)*

It is possible that the estimated impact of pre-teaching was diluted by there being a longer gap between pre-teaching and the maths lesson. See '[Perceived outcomes, mediators, and moderators](#)' for further discussion.

## Who received the pre-teaching

Year 1 and Year 2 teachers in participating schools were asked to select 12 pupils in their class who were then randomly allocated to the pre-teaching group or the comparison group. The trial used pupil-level randomisation to maximise the chance of detecting a difference in outcomes. Our expert advisors and teachers engaged in the scoping phase and pilot agreed that five to six pupils was the ideal size for small group support, and that more than six pupils was likely to become harder to manage logistically, partly due to increased differences in pupil levels and need, thereby reducing potential impact. This is also in line with evidence from the EEF on small group tuition (EEF, 2021). The decision to have six pupils was then based on the need to maximise statistical power, while managing teachers' abilities to deliver the sessions. Pupil selection guidance stated that the 12 pupils selected in each class should be those who stand to benefit the most from the additional targeted support in maths. The guidance gave the following examples to help teachers with their selection—pupils who are working below age-related expectations, experience maths anxiety, or lack confidence during the maths lesson, or tend to struggle with some key concepts and skills.

There was no strict inclusion or exclusion criteria for pupils in the pre-teaching group. The baseline survey asked teachers to rank the factors that were considered when selecting pupils, in order of importance. The results suggest that teachers mainly looked at age-related expectations (49% ranked this first) alongside their own 'subjective assessment of which pupils would benefit' (35% ranked this first). This aligns with the instructions that teachers were given. Overall, it is worth noting that there was a large amount of teacher discretion involved in the selection process.

The survey and interviews highlighted significant challenges with the selection process, which may have affected who received pre-teaching, and thus, moderated its effectiveness:

- First, some teachers felt that selecting pupils in the Summer Term for an Autumn Term approach was too early. They said a lot could change over the summer, particularly in this age range (they saw some pupils

regress and others progress) and would have preferred to nominate pupils in September based on an updated assessment.

- Second, and most importantly, some teachers said they did not have 12 pupils in a single class who they deemed eligible for the pre-teaching (e.g. working below expectations). This meant that some pupils were included in the pre-teaching group even though they were not those whom teachers believed could benefit the most from pre-teaching (i.e. they were working at or above expectations), resulting in a mixed-ability group. Simultaneously, some pupils who may have been ideal candidates for the pre-teaching group were instead randomised into the comparison group, and did not receive the pre-teaching.

Consequently, the results of the impact evaluation may not be a true reflection of the impact of pre-teaching when delivered to a targeted group of pupils in the absence of the selection constraints, which were implemented in the context of this trial (mainly to increase sample size and thereby maximise statistical power, with the aim of ensuring we were able to detect an effect, if it exists).

While teachers were not always satisfied with the groupings, the endline survey results suggest that schools complied with the core restriction of the trial to deliver pre-teaching to pupils allocated to the treatment group. This is based on the fact that the majority said that sessions were delivered to the group of pupils allocated to receive the sessions (97%) and that this group stayed fixed during the Autumn Term (98%). These results are based on the following questions in the endline survey: 'Have there been any changes in the group of pupils receiving the pre-teaching?' and 'Were any of the pupils who joined the pre-teaching group initially assigned to the comparison (no pre-teaching) group, in the list we shared with you?' (only teachers who selected: 'Yes, one or more pupils joined the group' to the first question were invited to respond to the second.)

### Relationship to main lesson

Teachers were asked to ensure that their pre-teaching sessions had a 'forward-looking focus'. This element is central to how we defined pre-teaching for the trial, as its effectiveness is believed to be linked to pupils previewing lesson content and warming up prior knowledge. The case studies show a mixed picture in terms of how far teachers stuck to this. On the one hand, interviewees described delivering a '*shorter and more simplified version of the lesson*' which consisted of using the same resources (e.g. slides), activities (e.g. manipulatives, style of questions), and language (e.g. key vocabulary) that would be used in the upcoming lesson or unit. However, some session leads also told us they had incorporated elements of catch-up (e.g. retrieval exercises, consolidating content from the previous lesson) in their sessions, where this was felt to be more beneficial for pupils' learning. We do not know how prevalent this was across the whole sample of schools, which limits our confidence that the impact evaluation results reflect the true effectiveness of pre-teaching when implemented as intended. There are several possible reasons for teachers incorporating elements of catch-up:

- **Usual practice.** Teachers said that they would typically start lessons with retrieval practice to embed learning and remind pupils of things they had already covered, and that small group programmes were almost always focused on post-teaching/addressing gaps based on teacher assessments during lessons (see section on 'Usual practice, differentiation, and comparison group activity' for more detail).
- **Interpretation of 'forward-looking'.** There is possible ambiguity about what counts as 'forward-looking'. For example, reminding pupils of a concept might still be important for enabling access to the lesson. Clearer examples/guidelines may be needed to ensure consistency.
- **Teacher use of guidance and understanding of guidelines.** Interviewees generally demonstrated a good understanding of pre-teaching that aligned with the guidance, however, in the endline survey just over half (55%) selected 'not very often' when asked how often they used the Teacher Guide. This may have affected fidelity. Reasons for not using the Guide were not having the time (26%) and not finding it useful (20%). Nevertheless, among those who said they used the Guide 'often'/'very often' (37%), the most common reason was for checking what you should/shouldn't do (59%). This suggests it was a useful tool for some in ensuring fidelity. There is potential for increasing the usefulness of the Guide as a practical tool, discussed further below.

## Quality

Teachers had autonomy over how they delivered the sessions. This section explores the quality of session delivery and how this varied:

- **Delivery style.** This was judged to be of high quality by observers. We observed teachers and teaching assistants use encouragement and praise to motivate pupils and acknowledge their efforts. They also employed a variety of strategies to keep pupils engaged, such as using manipulatives, varying their tone of voice, providing one-on-one support, and intentionally introducing mistakes as a teaching tool.
- **Session structure.** The observed sessions focused on a single concept (e.g. multiplication, shapes), in line with the guidance. However, the structure of these sessions varied. Some sessions began with whole-group work before moving to individual or paired practice, which allowed session leads to check for understanding and address misconceptions. Other sessions, particularly those that were shorter, remained in a whole-group format for the entire duration. These sessions offered fewer opportunities for independent practice and teacher feedback before the main lesson.
- **Activities.** According to the endline survey, explaining key concepts (64%), demonstrating a skill or strategy (57%), and group skill practice (53%) were the most frequently run activities. Teachers predominantly used storytelling and manipulatives (e.g. counters/chips, dice, dominos, shape blocks) in the sessions we observed, to build understanding of concepts. We observed the use of quizzes and songs/rhymes, which made the sessions interactive, but we did not see any games and this was a less common activity in the survey (12%). In the interviews, some teachers said that there was simply not enough time to introduce and play a game during the session.
- **Tailoring.** During the sessions we observed, teachers were attentive to pupils who needed extra support, for example, by repeating instructions and pairing up with pupils to complete the activity. However, teachers said having pupils with different levels of ability made it difficult to pitch the sessions at a suitable level for all pupils (as discussed in the section on '[Who received pre-teaching](#)'). In addition, some teachers said they had pupils in their group who needed one to one support due to behavioural issues or language and communication difficulties. As a result, they struggled to access the pre-teaching. This may have affected what different pupils got out of the pre-teach sessions, and therefore the effectiveness of the pre-teaching (this is discussed further in the section on '[Perceived outcomes and moderators](#)').

Several factors influenced the quality of sessions:

- **Role and level of experience of session leads at baseline.** Most, but not all, had prior experience of delivering pre-teaching to a small group (87%). There was a mixture of experience based on length of time in teaching, with just over half (57%) having ten years or more of experience, a quarter (25%) having between five and ten years of experience, and 18% having less than five years of experience. However, some of the project leads we interviewed said that teaching assistants, especially those with less experience, struggled to grasp the purpose of pre-teaching. This probably resulted in variation in quality of the sessions.
- **Length of sessions.** This may have impacted quality, with shorter sessions having less time for practice, play, and discussion.
- **Teacher Guide.** The Teacher Guide was designed to provide session leads with ideas for activities to use in the sessions, however, the survey results suggest that it was not used to its full potential. While interviewees found the example session plan helpful, less than a third (29%) of survey respondents reported using it for finding ideas for activities. Interviewees and survey respondents wanted to see more and a wider range of example activities and ideas covering different topics to support their planning.
- **Group composition.** The composition of the groups presented a significant challenge. Groups often included pupils with mixed abilities, making it difficult to effectively tailor the sessions to everyone's needs.

## Pupil responsiveness and engagement

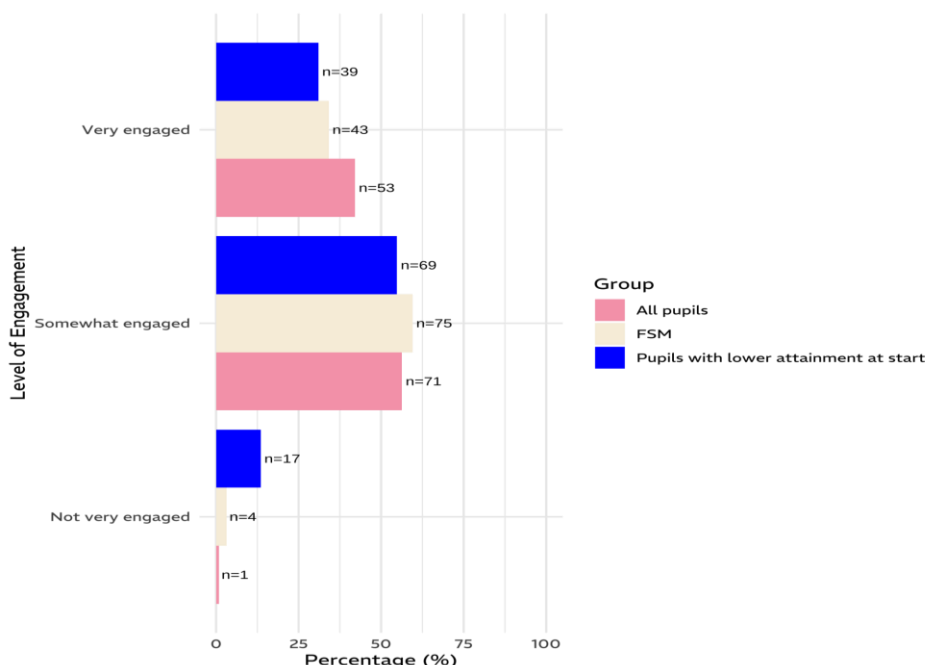
The findings suggest that, overall, there was a broad level of engagement with the pre-teaching. 98% of endline survey respondents agreed that pupils had been 'very engaged' (42%) or 'somewhat engaged' (56%). Only 2% of pupils were

considered by online survey respondents to be ‘not very engaged’. In the interviews, teachers said that engagement was supported by the following:

- **Perceptions.** Teachers said that framing the sessions as something ‘special’ or ‘extra’ helped with engagement. This was reflected in the focus groups where pupils described the sessions as ‘*early maths that no one else gets to do*’ (Child focus group).
- **Activities.** The use of manipulatives and storytelling made the sessions interactive and increased engagement. Pupils said they enjoyed the stories and were enthusiastic about using counters and shapes.
- **Session length.** The short length of the sessions helped, but teachers said there were still challenges with concentration and stamina, particularly with younger Year 1 pupils. This was also evident in the observations—session leads had to repeatedly use different techniques to bring back focus.
- **Timing.** Pupils were more alert and engaged in the morning. Approximately half of endline survey respondents (55%) delivered the sessions first thing in the morning, 29% in the morning lesson time, and 30% in the afternoon lesson time.

Engagement differed within the groups according to teachers. The endline survey results (see Figure 3 below) show that the proportion of teachers who considered pupils as ‘not very engaged’ was highest for low-attaining pupils (13%) compared to all pupils/FSM pupils. They were also less likely to be considered as ‘very engaged’. This was supported by the reflections of interviewees, who said that they found it challenging to engage pupils who found maths difficult *and* struggled with school/being in a classroom in general. In the observations, we saw this happen where some pupils sat further away from the group, did not actively participate, and showed signs of being bored and uninterested. They also found the subject matter difficult (e.g. were unable to answer questions correctly). Difficulties engaging some pupils may have undermined the effectiveness of the pre-teaching, which relies on pupils engaging with, and understanding, the material ahead of the main lesson.

Figure 3: Level of engagement with pre-teaching



## Usual practice, differentiation, and comparison group activity

Schools were asked about how they typically delivered targeted support in maths, to explore how this compared with pre-teaching, and understand what support pupils received during the trial period. Usual practice consisted of two main types targeted support:

- **Support during the lesson.** To help lower-achieving pupils to understand the lesson. This was the most common type of targeted support that treatment and comparison group pupils received during the trial (82% and 79%, respectively). It involved using year below resources, providing teaching assistant support within the classroom, and taking pupils out of the classroom during the lesson (although they tried to avoid this since it meant those pupils would miss out on conversations).
- **Post-teaching/catch-up sessions.** To fill gaps in knowledge and understanding. Around 40% of respondents said that pupils in the treatment group received this type of support, compared to 34% for pupils in the comparison group. This was felt to be distinct from pre-teaching, which teachers described as a more proactive approach to providing targeted support compared to catch-up programmes. Pre-teaching was not common practice in this form among case study schools, with the exception of 'soft starts'.<sup>18</sup> Some teachers/schools were interested in whether it would be more effective than what they normally do, and whether it could pre-empt other programmes, by addressing misconceptions beforehand.

The survey results show a low baseline prevalence of pre-teaching. At baseline, 6% of those whose school ran small group programmes said that this consisted of pre-teaching. At endline, out-of-class support before the lesson (i.e. pre-teaching) was minimal (4%) for comparison group pupils, which suggests that teachers complied with the requirement to deliver business as usual (not pre-teaching) to the comparison group. In addition, this shows that outside of the additional pre-teaching, treatment and comparison group, pupils likely received similar levels of support. However, in Pulse survey 3, 24% (n=28) of respondents selected 'pre-teaching' when asked what targeted support pupils in the comparison group received in maths during the trial period. While this would obviously represent a threat to the internal validity of the impact evaluation,<sup>19</sup> 18 of the 28 respondents who make up this 24% also responded to the subsequent endline survey, where they indicated that the pre-teaching was only delivered to pupils in the allocated group (and were therefore, classified as compliant for this). These contradictions highlight uncertainty about who received the pre-teaching and the level of non-compliance. Differences in responses might reflect differences in question wording (in Pulse survey 3, respondents were asked: 'What targeted support have pupils in the comparison group received in maths, if any?' of which pre-teaching was one of the options). Regardless, the remaining ten respondents (at ten different schools) who responded this way in Pulse survey 3 did *not* respond to the endline survey. Assuming they responded accurately, this indicates potential issues with compliance.

We do not have any information from these schools to explain this. However, we know from the interviews that there was some dissatisfaction with the allocations—for example, teachers felt that some of the lowest ability pupils were placed in the comparison group and were disappointed that they would not benefit from the pre-teaching. This concern may have led some teachers to introduce elements of pre-teaching to those pupils. As we did not collect pupil-level compliance data (e.g. total number of sessions, attendance at sessions), we are unable to report on exactly who received the pre-teaching and how much pre-teaching they received (dosage). Ultimately, it is worth recalling that the generalisability of these survey findings are limited by the low overall response rate in each case (34% and 46% for the baseline and endline surveys respectively; 25%, 26%, and 49%, respectively for all three pulse surveys), and that they should be interpreted with caution.

## Perceived outcomes, mediators, and moderators

### Perceived outcomes

#### *Confidence and anxiety*

According to the logic model, pupils who receive pre-teaching should experience reduced maths anxiety during the main lesson. The impact evaluation did not find a difference in this outcome between the comparison and treatment groups. However, the IPE found evidence that teachers believed pre-teaching can boost confidence in maths. Around 86% of endline survey respondents said they noticed a positive difference in pupil confidence over the course of implementation, though the survey did not ask teachers specifically about maths anxiety. This was supported by the findings from the interviews, as shown by the quote below:

<sup>18</sup> Soft starts refer to a gentle, flexible way to begin the school day, designed to help pupils transition smoothly into learning.

<sup>19</sup> It is worth noting that non-compliance in the form of pupils allocated to the control condition receiving the pre-teaching would lead to an *underestimation* of any positive impact of pre-teaching on measured outcomes, rather than an overestimation.

*It's been really useful for the children to have that small group for them to feel more confident to come into the classroom and access it more independently. (Teacher interview)*

Teachers reported that as a result of increased confidence, pupils persevered with independent tasks, were more willing to use the resources/manipulatives and actively participated in class discussions. Teachers said that being exposed to the work beforehand gave pupils the motivation to 'have a go' during the main lesson, and the tools to engage with what was being asked of them. This suggests that the following mechanisms are important:

- **Strong bridge-building between pre-teach and lesson content.** This includes during the main maths lesson. In the interviews teachers said they would remind pupils in the pre-teaching group that they had already practiced a task and would refer back to what they did in the small group. In the focus groups, pupils who agreed that the sessions were helping them with maths said that this was because they were doing the same thing that they would be doing in the lesson.

*I was able to say to them during that group that I can't wait to see what you do tomorrow in maths. So then within the math lesson, I would remind them, do you remember yesterday during that group we did? (Teacher interview)*

- **Explicitly teaching pupils self-regulation strategies (in the pre-teach sessions).** The logic model links pupils being better able to self-regulate (manage their motivation, thoughts, and behaviour to set goals, monitor working, and reflect and review progress) to lower anxiety. Examples of this from the observations included teachers modelling how to answer a question, framing mistakes positively (as something to learn from), and teaching pupils how to work through them (telling them that this was what makes a good mathematician).

Nevertheless, confidence gains were also attributed to the small group aspect of pre-teaching due to this being a 'low-stakes' environment where pupils could access additional support from a teacher/teacher assistant.

*She wouldn't normally put her hand up to answer questions whereas in this small group of five children she had the confidence to give answers and I think that really helped her. (Teacher interview)*

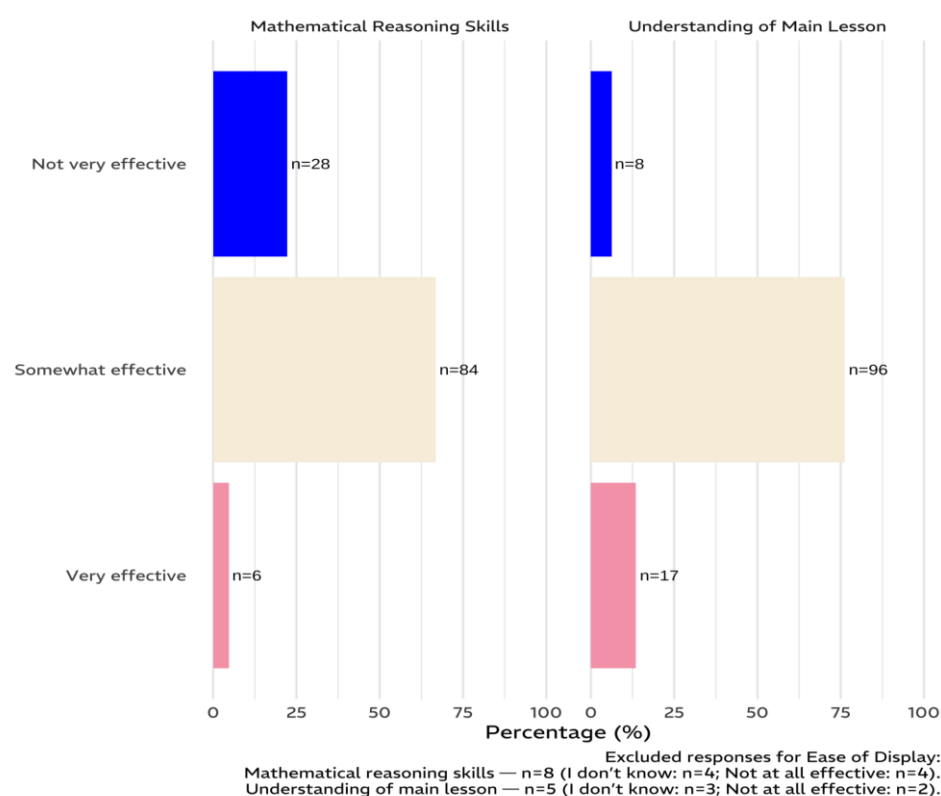
The logic model suggests that pre-teaching can also boost confidence by changing how the wider class perceives members of the group. There was some evidence of this, for example, where teachers reported that the pre-teaching group was viewed positively by other pupils, who sometimes asked to join in. This was reflected in the focus groups with pupils where pupils described the group as something 'extra' that only they got, which their classmates wanted to join. The way teachers framed the group, such as calling it a 'maths club', contributed to this perception. Over time, it is possible that this also shifts attitudes about who is 'good' or 'bad' at maths, further reinforcing confidence of pupils in the group. However, the logic model also acknowledges the risk of pupils in the pre-teaching group feeling labelled and stigmatised. This was observed by some teachers early on in the trial period, who were concerned that being picked for the pre-teaching group made certain pupils even less confident as they felt they were being singled out due to their low skill level. To mitigate this, they reframed the group as a way to 'stretch' pupils, which provided reassurance. This highlights the importance of framing in minimising the chance of backfire effects, although it is unclear how successful this was in practice.

#### *Understanding of lesson content and mathematical reasoning*

The logic model suggests that pre-teaching can help pupils to better understand the mathematical concepts and language covered in the upcoming main lesson, by helping to retain topic knowledge and master key concepts and vocabulary. It is hypothesised that in the long term, this will have the result of improving their mathematical reasoning, which is the primary outcome measure in the impact evaluation.

The survey results show that pre-teaching is perceived to improve pupils' understanding of lesson content and, to a lesser extent, mathematical reasoning skills. At endline, most respondents (90%) said that pre-teaching was effective at boosting their pupils' understanding of the main lesson. Fewer teachers (70%) said that pre-teaching was effective at improving mathematical reasoning. In the interviews, while they were positive about the confidence gains, teachers found it difficult to comment on whether skills had improved, either because it was too early or because they felt this needed to be verified by assessment data. Data from the focus groups was limited in general, but some pupils said that the pre-teaching sessions made them 'smarter'.

Figure 4: Perceived effectiveness of pre-teaching



Teachers who said they thought pre-teaching was effective at improving mathematical reasoning skills were asked why they thought this to be the case. The results are shown in Table 24 below.

Table 24: Reasons why pre-teaching was effective at improving mathematical reasoning

Benefit	n	%
They boosted their confidence	64	51
They provided additional teaching time	43	34
They improved the retention of key concepts	29	23
They improved their ability to use basic mathematical vocabulary	14	11
They made pupils more autonomous during the lesson	14	11

Confidence was the most important driver (51%; Table 24). In the interviews, teachers linked this to increased participation in maths lessons, which they saw as having a knock-on effect on skills (Table 24). Around 23% of respondents selected 'they improved their retention of key concepts' and 11% of respondents selected 'they improved their use of basic mathematical vocabulary' (Table 24). These are hypothesised to be key mediators related to an improved understanding of lesson content in the logic model. The following factors may explain why these mechanisms were weaker than expected, and help to interpret the results of the impact evaluation, which found a very small (negligible) positive effect of the pre-teaching on pupils' mathematical reasoning skills.

### The pre-teaching was not intense or consistent enough

There were several factors related to how the pre-teaching was implemented that are likely to have contributed to the lack of effect found in the impact evaluation.

First, the implementation period itself was short by design (delivered over one term), and staffing and timetabling challenges were common. This may have affected how much of the pre-teaching pupils received and the consistency of delivery. In the

interviews teachers said they thought pre-teaching would be more impactful if delivered more often, but that this wasn't feasible. Indeed, the findings suggest that implementation was less feasible than expected based on feedback from teachers and experts gathered during the scoping and pilot phase. This discrepancy may be due to a combination of factors. The teachers and schools we heard from during scoping/the pilot may have been particularly motivated and enthusiastic, or they may have underestimated the challenges associated with delivering pre-teaching sessions alongside many other competing demands.

Second, we know that sometimes the sessions were not entirely 'forward-looking'. They were also short in length, and there is some doubt around whether the shortest sessions (less than ten minutes) include enough time for pupils to discuss uncertainties and unpick misconceptions with the teacher before they arise in the main lesson.

Challenges related to time and staff availability are likely to be barriers to teachers/schools introducing this into their everyday practice in the longer term. However, training and/or guidance for session leads could help to support the implementation of forward-looking pre-teaching.

### The groups were too broad and not targeted enough

The design of the trial was underpinned by the assumption that there would be 12 pupils in each class who would be a good fit for pre-teaching. This was motivated by scoping phase findings and our expert advisors' input, as well as existing evidence on small group support, balanced out against sample size and statistical power requirements. As discussed earlier (see '[Who received the pre-teaching](#)' section) this did not prove to be systematically true—some teachers said they struggled to put forward 12 pupils who were working at a similar level (below expectations). In turn, this may have led to certain groups in some schools being too broad and not targeted enough. Having a mixture of abilities in the group was the most selected reason for pre-teaching not being effective at improving mathematical reasoning (63% selected, see Table 25 below). Interviewees who had pupils of varying abilities in the group said that it was difficult to make the session suitable for all pupils. As a result, they said that the pre-teaching was not as targeted as other small group programmes. These barriers can be seen as related to the trial requirements. Schools who choose to implement pre-teaching as part of usual practice will be able to create more targeted groups of pupils.

Table 25: Reasons why pre-teaching was not effective at improving mathematical reasoning

Challenge <sup>a</sup>	n	% total sample	% respondents shown item
There were large disparities in prior maths understanding in the group	20	16	63
The pre-teaching sessions were not regular enough	8	6	25
There were too many pupils in the group	6	5	19
Poor pupil attendance	4	3	13
The length of the project was not sufficient	3	2	9
Pupils were not engaged with the pre-teaching sessions	3	2	9

<sup>a</sup> Please note that this question was only shown to n = 32 respondents who state that pre-teaching was 'Not very effective' or 'Not at all' effective in improving pupils' mathematical reasoning skills. This data was gathered in the question prior, phrased as: 'On average, how effective would you say the pre-teaching sessions have been at improving your pupils' mathematical reasoning skills?' Respondents could select multiple reasons and thus raw n values in Table 25 do not add up to n = 32.

Other possible factors that may have moderated the effects of pre-teaching are discussed further below under 'Moderators'.

### Moderators

#### Pupil characteristics

Teachers felt that pupil needs, level of engagement with school, and attendance affected what pupils got out of the pre-teaching.

- **Needs.** Teachers described difficulties meeting the needs of pupils in the group where there were large discrepancies. Some felt that there were pupils, such as those with additional needs (e.g. Special Educational Needs and Disabilities [SEND]) who needed targeted one-on-one support to get the most out of pre-teaching.
- **Engagement.** As reported in the section on '[Engagement](#)', teachers considered their groups to be engaged overall, but they said that focus and participation varied between pupils. Where this was lower, teachers were concerned that the pre-teaching was less beneficial. This also came across in the focus groups where there was a clear split between pupils who thought the session helped them, and pupils who did not (the latter being those who said that they did not like maths and they would rather be outside playing).
- **Attendance.** The fact that the sessions followed on from each other meant that pupils with low attendance were disadvantaged due to having gaps in their knowledge. 25% said that attendance was a key challenge in the endline survey.

#### *Teacher characteristics*

Staff experience was cited as one of the top factors that helped with implementation (57%) but delivery was more challenging for less experienced teachers/teaching assistants. Some felt that pre-teaching had been more effective where the class teacher was delivering.

#### *School characteristics*

- **Implementation timing.** Schools began delivering at different times, due to delays completing assessments. As reported in the impact evaluation, the regression analysis suggests that pupils who were in schools that delivered the pre-teaching 'on time' scored slightly higher than pupils in the 'late' group in the endline assessment. We included both 'on time' and 'late' schools in the case study sample but did not find any noticeable patterns to explain this.
- **School size.** Multi-form entry schools faced particular challenges in timetabling, as they needed to schedule two pre-teaching sessions per class each week to meet the trial requirements. Finding enough time slots was difficult, which could act as a barrier to consistent delivery.
- **Physical environment.** Around 45% of respondents to the endline survey cited availability of physical space as an enabling factor. Teachers said they had 'pockets' of spaces available for small group work.

#### *Implementation factors*

Most of these factors are discussed elsewhere in the report, so we only include a high-level summary here:

- **Group composition.** There was the view that the groups were not targeted enough and too heterogeneous in terms of ability.
- **Consistency of delivery.** Teachers were required to deliver the sessions at least twice a week for one term. The endline survey suggests that this was largely adhered to, however, there were challenges with timetabling and staffing, which may have affected how much pupils received.
- **Duration.** A quarter (25%) of sessions had been shorter than the recommended length (less than ten minutes). Teachers reported difficulty keeping Year 1 pupils engaged for longer periods due to limited attention spans. There is doubt whether sessions shorter than ten minutes provided enough time to meaningfully introduce concepts and address misconceptions.
- **Teacher guidance.** Just over half of respondents said that they did not use the Teacher Guide often during the trial. This lack of engagement with the Guide could have negatively affected fidelity and quality. Some teachers also expressed a desire for more resources and activity ideas within the guidance.
- **Relationship to main lesson.** Teachers said that explicitly linking the pre-teach and main lesson, by reminding pupils that they had already practiced the tasks, helped pupils make connections and boosted their confidence and engagement in the main lesson. This 'bridge-building' was identified as a key mechanism for positive outcomes.

We asked teachers how they felt about continuing to implement pre-teaching beyond the trial. Overall, teachers were keen to continue. At endline, 83% of respondents said that they would feel confident doing so and 84% of respondents said that they would recommend pre-teaching to other teachers. They highlighted the following as important if adopting pre-teaching outside of the trial:

- Teachers wanted to be able to make changes to groups to ensure they were as targeted as possible to address the challenges described above.
- Teachers wished to have control over when and how they assess pupils to determine if pre-teaching is working. They found the trial's assessment administration to be challenging and time-consuming (discussed further below).
- Teachers would make changes to ensure delivery could be sustained in the long term, by modifying the frequency or duration of sessions to better fit daily school routines.
- Teachers wanted to see a wider range of example activities aligned with specific topics/concepts to inform their planning.

## Outcome data collection

In each school pupils who were put forward for the pre-teaching were required to complete a baseline and the endline assessment of the primary outcome measure (QRT) as well as CMAS-UK at endline, to measure the secondary outcome of pupils' anxiety related to maths. The assessments were administered by teachers, which has potential disadvantages related to burden, quality and consistency, and bias. The IPE therefore, explored how well teachers understood and complied with the outcome test administration guidelines.

The feedback we received suggests that teachers experienced significant difficulties administering the assessments, despite being given extensive guidance ahead of time on how to deliver them to pupils. Around 65% of endline survey respondents said that they found this 'difficult'/'very difficult'. The assessments were time-consuming to administer, particularly to Year 1 pupils, which were completed on a one to one basis to ensure those pupils were adequately supported during the process. Around 55% said that time was one of the main challenges they faced. There were delays to schools completing and submitting assessments, which may reflect the time required to complete them, and contributed to the decision to grant an extension to those schools unable to start delivering the pre-teaching sessions after 30 September 2024 (creating the 'on time' and 'late' school subgroups documented above).

The guidelines themselves were generally positively received. Approximately half (52%) agreed that the instructions met their needs, while 21% selected 'neither agree nor disagree'. Around 13% selected 'disagree'/'strongly disagree'. In the interviews, teachers who experienced difficulties with the guidelines said that they had to read over the guidance multiple times, and would have liked an opportunity to ask questions. They also thought the guidance could have been shorter and less wordy. This highlights some areas for improving the guidance to make it easier for teachers to digest, although the majority of teachers still found the guidance to be appropriate, and ensuring that the necessary information was there for those teachers who needed it may outweigh the desire for more streamlined guidance for those who did not.

We also received feedback from teachers on the outcome measures used to assess pupils, in particular the QRT for mathematical reasoning. There were concerns that the test would not reflect the true ability of pupils—some questions were seen as being too difficult and not relevant to the curriculum or the scheme of work schools were following. However, our findings suggest that while some teachers judged the QRT to be challenging, it was able to effectively measure a range of abilities, and was not so difficult that it created a floor effect. This is evident in the approximately normal distribution of scores, which can be seen in Appendix F (i.e. scores are not bunched around a '0' or low mark). It is worth noting that the chosen measure was intended to be 'topic agnostic', and focused on mathematical reasoning, which means that technically curriculum coverage should not have been an issue. The test developers also leveraged their expertise to ensure the test was age-appropriate. However, the views of teachers are important here, especially as they affect how they would have administered the test in practice. This does not suggest that the measure for mathematical reasoning was necessarily

flawed, but points to the importance of factoring in teachers' views on the outcome measure (especially if they administer the assessments themselves) during the design, with the lack of a formal process to do so in this evaluation being a limitation caused by the resource and timeline constraints faced.

## Conclusion

Table 26: Key conclusions

Key conclusions	
1.	There is no evidence that pupils allocated to receive the One Step Ahead maths pre-teaching made any additional progress in their mathematical reasoning skills, on average, compared to pupils not allocated to receive the pre-teaching. This result has a moderate to high security rating.
2.	Among children eligible for free school meals (FSM), the best estimate of pre-teaching's impact is a small negative effect on pupils' mathematical reasoning skills. These results may have lower security than the overall findings because they relate only to a subsample of pupils.
3.	Teachers faced challenges with timetabling and staff availability, with many reporting difficulty fitting sessions into their daily schedule. This led to inconsistent delivery and sessions that were often shorter than the recommended 10–15 minutes. School size may have been a factor here, with some teachers in multi-form entry schools noting particular challenges in timetabling.
4.	Teachers perceived pre-teaching to have a positive effect on pupils' confidence and engagement, teachers attributed this to small groups' low stakes environment, teachers building explicit links between the pre-teaching and the main lesson ('bridge building'), and changes in how the wider class perceived the pupils in the group. These perceptions from teachers were not reflected in the impact evaluation outcomes of maths attainment and maths related anxiety.
5.	The pupil selection process also posed challenges, and randomisation sometimes led to pupils who teachers felt would not benefit most from pre-teaching (i.e. they were working at or above expectations) being included in the pre-teaching group. This meant that some pupils were included in the pre-teaching group even though they were not the intended target, resulting in a mixed-ability group. These factors made it harder to tailor sessions effectively.

## Impact evaluation and IPE integration

### Evidence to support the logic model

The trial's findings offer a nuanced perspective on the effectiveness of the One Step Ahead pre-teaching, providing partial but valuable support for its underlying logic model. While the impact evaluation suggests that there was no effect on the primary outcome of pupil mathematical reasoning skills or its secondary outcome (maths anxiety), the IPE provides helpful context and evidence for several of the causal links hypothesised within the original logic model.

There is some evidence in support of the logic model in the IPE's qualitative and quantitative findings on pre-teaching's perceived impact and mediating mechanisms. The model posited that pre-teaching would boost pupil confidence, improve retention of topic knowledge, and foster more active engagement in the main maths lesson. These assumptions were largely supported by teachers and pupils in the IPE analysis. A significant majority of endline survey respondents (86%) reported a positive change in pupil confidence. This was reported as being a direct result of the pre-teaching sessions being a low-stakes environment where pupils could practice key concepts and vocabulary without the pressure of a full class setting. The qualitative data also provided evidence for the 'bridge-building' mechanism, with teachers describing how they explicitly linked pre-teaching content to the main lesson. This practice, identified by our expert advisors as a key 'active ingredient', was perceived by some teachers to be effective in boosting pupil self-efficacy, and pupils were perceived by these teachers to be more able to actively participate in class discussions and activities because they had already had a 'preview' of the material. Among teachers who responded to survey questions on the reason(s) for why pre-teaching was (not) effective at improving pupils' mathematical reasoning 23% selected 'they improved their retention of key concepts' and 11% selected 'they improved their use of basic mathematical vocabulary'. These are hypothesised to be key mediators related to an improved understanding of lesson content in the logic model.

The discrepancy between these positive perceived outcomes and the null findings of the impact evaluation is important to note. It is possible that teachers may have been overly positive or optimistic in their assessment of how beneficial the pre-teaching sessions were, at least in relation to the outcomes measured in this evaluation. It is worth noting also that some teachers did raise doubts about whether the shorter pre-teaching sessions (e.g. those that were ten minutes rather than 15 minutes) were enough time for pupils to discuss uncertainties and unpick misconceptions with the teacher before they arose in the main lesson. Relatedly, the discrepancy between the often, positive findings from the IPE and the null findings from the impact evaluation, are likely to be explained in part by contextual moderators and implementation challenges. The

logic model seems to have correctly identified factors such as staff time and resources, group composition, and the consistency of delivery as potential moderating factors, and the IPE found that these had a significant influence on fidelity and dosage. Teachers faced substantial challenges with timetabling, with 72% of survey respondents citing this as a barrier. This a key factor in ensuring consistent implementation, and this inconsistency, combined with the trial's short, one-term duration, likely diluted any potential treatment effect. Furthermore, the core assumption that teachers could select a targeted group of 12 pupils who would all benefit from pre-teaching was not fully met. Teachers reported struggling to identify this many pupils with appropriate and similar needs, resulting in mixed-ability groups that were difficult to tailor to effectively. This heterogeneity within groups could have prevented pre-teaching from delivering the necessary intensity of support to the lowest-attaining pupils.

The trial's findings highlight that while the causal logic has some supporting evidence in the IPE findings, the effectiveness of pre-teaching is likely contingent on a context where implementation can be delivered with an appropriate level of fidelity and dosage. For example, the impact evaluations subgroup analysis of 'on time' versus 'late' schools revealed a small but noteworthy difference in pupil outcomes, providing some quantitative support for the IPE finding that consistency of delivery is crucial. Further analysis shows this effect was largely driven by Year 1 pupils, for which 'on time' schools consistently scored higher on the QRT at both baseline and (in particular) endline compared to 'late' schools. In contrast, for Year 2 pupils, the differences between 'on time' and 'late' schools were negligible. This suggests that implementation consistency may be particularly critical for younger pupils, or perhaps that the direct one to one administration of Year 1 assessments by teachers (which was not the case for Year 2 assessments) may have magnified the effect of consistent delivery. The evaluation also identified a significant limitation with the chosen outcome measure. Feedback from some teachers suggested that the QRT did not feel relevant to their curriculum and was perceived as too difficult for some pupils, potentially influencing pupil performance and contributing to the lack of a measurable impact.

In conclusion, while this evaluation finds some support for the core premise of pre-teaching (i.e. that a small, targeted approach can boost confidence and improve understanding), it also finds that achieving a meaningful impact at scale is challenging, in significant part due to the implementation challenges described above.

## Interpretation

The trial's findings on the perceived positive impact on pupil confidence and engagement align with prior evidence from the scoping phase and other studies. The theory that pre-teaching could act as a 'booster' for confidence and understanding was supported by teachers and pupils, reinforcing the core causal logic outlined in the logic model. The study's specific contribution lies in testing this hypothesis within a 'Teacher Choice' model, which aimed to fill a gap in evidence regarding the best way to provide out-of-class maths support for young pupils. While some research suggests that the frequency and 'intensity' of pre-teaching correlate with attainment outcomes (van den Berg *et al.*, 2017), the overall low prevalence of the practice in UK schools (24% of surveyed teachers reporting they do so regularly or occasionally) meant there was a significant need for a trial to generate evidence on its impact. The null result from the impact evaluation, despite the positive findings from the IPE, points to the crucial role of implementation fidelity and dosage—factors that prior evidence had already flagged as key for effectiveness. Therefore, this finding should be interpreted as the impact of pre-teaching delivered under the specific, real-world implementation constraints and short duration of this 'Teacher Choice' trial, rather than a definitive statement on the effectiveness of all pre-teaching models.

The core hypothesis, supported by existing evidence on targeted support and our own scoping research, was that pre-teaching would act as a 'booster' to improve confidence, understanding, and ultimately, mathematical reasoning. The IPE findings provide some evidence for the initial stages of this causal chain. Teachers and pupils reported benefits of pre-teaching, particularly in boosting confidence and increased engagement in the main lesson. A key mechanism for this, confirmed by interviews, was 'bridge-building'—explicitly linking the pre-teaching sessions to the main lesson. This finding supports a central 'active ingredient' of pre-teaching and provides a clear, actionable insight for teachers. However, these positive proximal effects did not translate into measurable distal outcomes in the impact evaluation. One of the reasons for this appears to be that pre-teaching's fidelity and dosage were compromised by real-world challenges, along with a primary outcome measure (the QRT), which some teachers felt was unsuitable for Year 1 pupils in particular.

The trial's findings have significant implications for whether pre-teaching can be effectively replicated using the same resources and in normal practice conditions. While the Teacher Choice was designed to be low cost and low burden, the IPE revealed that even these minimal requirements were difficult for some schools to sustain. Some teachers cited a lack of staff time and resource constraints as the biggest barriers to consistent delivery. The short length of sessions (sometimes less than ten minutes) and the difficulty in timetabling them consistently suggest that the 'at least twice a week for one term' dosage was not always met. Interviews with teachers from the IPE highlighted that increasing the length of the sessions to provide a higher dosage would have been difficult to fit into their daily schedules. This suggests that a simple increase in dosage may not be a feasible route to improving impact and that the challenge lies in protecting the time and resources for the current model. However, the null impact finding in this evaluation should not preclude exploration of alternative, more intensive, or longer-duration pre-teaching models that may require greater resourcing and a different implementation approach. The impact evaluation's subgroup analysis, which found that pupils in 'on time' schools (those who started pre-teaching without delay) performed slightly better on the QRT than those in 'late' schools, provides quantitative support for the IPE's qualitative finding that consistency of delivery is crucial. These findings suggest that for pre-teaching to be effective, schools need to be able to protect the time and resources needed for consistent delivery.

The results of this study contribute to the existing evidence base by highlighting the importance of implementation fidelity, even for non-prescriptive, 'Teacher Choices'. While prior research on targeted maths support has shown positive effects (Sahasranaman *et al.*, 2025), this study uniquely focuses on the specific pedagogical choice of 'pre-teaching'. It adds to the body of evidence by demonstrating that while the a priori evidence for pre-teaching is strong, a key challenge is the translation of this theory into consistent practice in a busy, resource-constrained school environment.

To maximise the utility of the results, the key message is framed around a 'theory of change' informed by the IPE findings, moving beyond a binary 'yes' or 'no' on effectiveness and acknowledging the complexity of translating research into consistent practice. The implementation lessons derived from this trial can serve as a guide for teacher decision-making, explaining that while pre-teaching holds promise for boosting pupil confidence and engagement (as corroborated by the IPE), its success may be contingent on specific implementation conditions that were likely often unmet in this evaluation. This framing would empower teachers to make informed decisions about whether and how to implement pre-teaching in their own contexts. For example, a teacher could be advised that if their school can dedicate a consistent, protected time slot for these sessions, and if they have a clearly defined group of pupils with similar needs, pre-teaching is a promising strategy. Conversely, if these conditions cannot be met, they may consider an alternative approach.

We contend that the IPE findings suggest that the core issue experienced by teachers was not that the prescribed length or quantity of sessions (10–15 minutes, twice a week) was theoretically insufficient, but rather that the lack of quality and consistency of implementation, manifesting in poor adherence to the intended dosage and deviation from the forward-looking delivery model, likely diluted the potential beneficial effects of pre-teaching. Additionally, pre-teaching was designed as a low cost, low burden strategy to minimise displacement of other activities, with the pre-teaching sessions being intentionally brief (10–15 minutes). In sum, we believe that challenges experienced by teachers were less about the inherent cost of displacement and more about the systemic difficulty for teachers of protecting any time at all for new initiatives in a packed school schedule.

## Limitations and lessons learned

The evaluation, while providing valuable insights, was subject to several limitations and challenges in its design and implementation that could have impacted the results presented here. Reflecting on these is crucial for drawing accurate conclusions and informing future research in this area.

One of the most significant challenges was the dropout rate for schools. While the attrition rate between randomisation and endline (4 of 75 schools, 5.3%) was similar to the rate (6%) assumed in the original power calculations, a large number (25) of additional schools withdrew from the evaluation between signing up and randomisation itself. This, in turn, affected the trial's ability to detect a significant effect. While the final sample size was salvaged by the over-recruitment of multi-form entry schools, the high-attrition rate introduces a potential for bias. It is possible that schools, which dropped out, differed

systematically from those which remained, perhaps in their level of commitment or resources, which could threaten the generalisability of the findings. A potential lesson for future trials is to factor in a more conservative attrition rate.

The reliance on school staff for performing data collection also presented a number of challenges. The IPE found that teachers experienced significant difficulties in administering the outcome assessments, with 65% of endline survey respondents finding the process ‘difficult/very difficult’. This was primarily due to the assessments being time-consuming, particularly the one to one administration required for Year 1 pupils. This burden likely contributed to the high number of missing test results (423 pupils excluded from the final analysis, an attrition rate of 16.2% of pupils for the primary outcome between baseline and endline), and the delays in data submission. In future research, alternative, less burdensome assessment methods or the use of external assessors could be considered to improve data quality and minimise teacher workload.

Furthermore, the study’s data collection methods, particularly the use of pulse surveys, had limitations that impacted our understanding of key implementation factors. The low response rate to all three pulse surveys (only 18% of teachers responded to all three) makes these findings susceptible to bias, endorses our framing of them as ‘temperature checks’ rather than as a reliable method for collecting consistent monitoring data. The simplicity required to encourage completion also meant that crucial information, such as the exact dosage (number of sessions) and pupil attendance, could not be captured. This lack of precise dosage data is a substantial limitation, as we cannot definitively state whether the lack of impact was due to the pre-teaching itself, or because pupils did not receive enough pre-teaching.

A further challenge was the potential for non-compliance. While the majority of teachers reported delivering pre-teaching only to the allocated group, there was some anecdotal evidence of non-compliance in the comparison group. In Pulse survey 3, a small number of teachers (24%) reported that pupils in the comparison group had received pre-teaching. This may have been driven by a misunderstanding of the survey question, or by teacher dissatisfaction with the random allocation (particularly when low-ability pupils were assigned to the comparison group). This situation may have introduced an ethical challenge for some teachers, forcing them to withhold perceived beneficial support from pupils they had identified as likely to need it, thereby increasing the internal motivation for potential non-compliance. This raises concerns about contamination, where the comparison group inadvertently receives elements of the approach, and it highlights a limitation in the data, as we were unable to collect pupil-level compliance data to confirm who received pre-teaching and for how long. Furthermore, the trial faced the risk that teachers might subconsciously modify their pedagogy in the main lesson (e.g. teaching a topic in a different way), or attempt to compensate for the imbalance by dedicating extra in-class attention to comparison group pupils who were perceived to be struggling, threatening the genuine ‘business as usual’ state of the control condition. However, it is worth recalling the discrepancy between Pulse survey 3 and the subsequent endline teacher survey, which suggested that schools complied with the core restriction of the trial to deliver pre-teaching to pupils allocated to the treatment group. This is based on the fact that the majority of respondents to the endline survey said that sessions were delivered to the group of pupils allocated to receive the sessions (97%) and that this group stayed fixed during the Autumn Term (98%).

In a similar vein, anecdotally we suspect that a large proportion of pupils who received pre-teaching during the trial did so from a teaching assistant—rather than their class teacher—on at least one occasion. While we have no equivalent endline figure, within baseline teacher surveys, 52% of respondents expected at least one teaching assistant to be involved in delivering pre-teaching in their school. While our initial scoping phase and Teacher Guide pilot found that giving schools flexibility on this (i.e. allowing teachers to rely on teaching assistant support if and when needed) was seen as positive (provided that the Guide contained enough guidance), the EEF’s evidence-based guidance on the deployment of teaching assistants clearly lays out the challenges of using teaching assistants effectively (EEF, 2025a). Specifically, the EEF recommends that teaching assistants are not used to replace teachers, and that pupils who struggle most should spend more time with their teacher (not their teaching assistant) as compared to their peers. The EEF’s guidance (EEF, 2025a) does not rule out using teaching assistants to deliver well-chosen, evidence-based, structured programmes where appropriate, but it may be that this particular type of approach is not appropriate for teaching assistants, especially in resource-constrained schools, where they may not receive the support needed to deliver it well.

Another threat to the study's internal validity and the generalisability of its findings stems from the pupil selection process and the resulting group composition. We note that some teachers struggled to identify 12 pupils who were suitable for the pre-teaching (as defined by the trial guidance), leading to the inclusion of pupils with a mixed range of abilities, some of whom were not the intended target.<sup>20</sup> This somewhat contradicts the spirit of the trial's design, which was based on the premise that pre-teaching would be delivered to a highly-targeted group of pupils working below age-related expectations. The qualitative data from the IPE confirms that this was a moderating factor, as it made sessions difficult to tailor and may have diluted pre-teaching's effect. Specifically, having a mixture of abilities in the group was the most selected reason for pre-teaching not being effective at improving mathematical reasoning (16% selected). Interviewee teachers who had pupils of varying abilities in the group said that it was difficult to make the session suitable for all pupils. As a result, they said that the pre-teaching was not as targeted as other small group approaches. Given the initial project scoping phase (conducted from October 2022 to January 2023) suggested that the nomination of 12 'eligible' pupils per class was feasible in this context, this suggests that future evaluations may wish to demand a higher standard of evidence for supporting key assumptions in trial design.

Similarly, the fact that some teachers incorporated 'catch-up' elements into their 'pre-teaching' sessions—a deviation from the intended forward-looking design—is a source of potential contamination and limits our confidence that the impact evaluation results truly reflect the intended Teacher Choice. The evaluation also highlighted a tension within the 'Teacher Choice' model: the balance between providing guidance and allowing teacher autonomy. While most teachers demonstrated a good understanding of the pre-teaching concept, over half of IPE survey respondents (55%) reported not using the provided Teacher Guide often. Reasons for this varied, but included a lack of time (26%) and a belief that the Guide was not useful (20%). Less than a third (29%) of survey respondents reported using it for finding ideas for activities, while interviewees and survey respondents wanted to see more and a wider range of example activities and ideas covering different topics to support their planning. Those who did use it often did so to check what they 'should/shouldn't do', indicating its value as a tool for ensuring fidelity. This finding suggests that while some teachers may have comfortably implemented the pre-teaching on their own, a more accessible or integrated guide could improve fidelity for others. Future efforts could explore shorter, more interactive, or digitally integrated formats that are less burdensome and more engaging for teachers, without sacrificing the professional judgement that is central to the Teacher Choice approach.

Finally, while the IPE provided valuable context, it also faced some limitations, particularly with the pupil focus groups. The challenges in eliciting responses from young pupils, especially those in Year 1, limited the depth of the data collected. This means that while we have some broad insights into engagement, the findings from this data source are not as prominent as those from the teacher surveys and interviews. It is possible that (with less resource and timeline constraint) this data collection approach being piloted may have allowed it to be refined to improve its effectiveness.

## Future research and publications

This evaluation provides a strong foundation for future research. While the primary impact evaluation finding showed no overall impact, the nuanced insights from the IPE highlight several key questions that require further investigation. One potential area for future research is to determine if a higher dosage of pre-teaching—for example, more frequent or longer sessions—can lead to a statistically significant impact on pupil outcomes. Research could also explore whether more prescriptive pupil selection criteria (and/or different groups sizes for the pre-teaching group), could lead to more targeted and effective approaches. Further research could focus on how to overcome the contextual barriers of timetabling and resource allocation. Additionally, future trials could benefit from exploring alternative outcome measures that are perceived as more relevant and valid by teachers, potentially in addition to the QRT, to ensure the full spectrum of the approach's effects are captured. Further alternatives and/or additions to the research include a longer delivery period (e.g. a full school year) or a different delivery period (e.g. Spring Term rather than Autumn Term), pre-teaching with other year groups, and offering more prescriptive pre-teaching guidance—noting that each of these potential research avenues have their own complications.

---

<sup>20</sup> It is important to consider that reducing the pre-teaching and comparison group sizes from six to four or three would have (due to requirements for statistical power) meant a commensurate increase in the number of schools to be recruited by 33% or 50%, respectively.

Specifically in relation to this latter suggestion, we believe that the Teacher Choice approach (with its emphasis on flexible guidance rather than strict fidelity rules) proved to be both a strength and a weakness of this trial. The flexibility made the Teacher Choice appealing and feasible for schools, however, it also introduced significant variation in implementation, which potentially diluted the overall impact. This points to a key challenge for Teacher Choice research: finding the right balance between flexibility (which encourages adoption) and fidelity (which is necessary for effectiveness). Future research in this area could explicitly investigate this trade-off, with a more prescriptive guide with fewer suggested activity options potentially leading to higher implementation fidelity. Relatedly, the suitability of the pre-teaching to be delivered by teaching assistants (rather than by class teachers), and the conditions under which this varies, is another potential avenue of research.

There are no plans for further publications, though the detailed data and analysis will be made available in the EEF data archive for other researchers to use. This will allow the research community to perform additional and exploratory analyses, such as exploring alternative statistical models or combining the data with other similar studies to investigate the impact of pre-teaching in different contexts and with a larger sample.

## References

- Austin, P.C. (2009) 'Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research'. *Communications in Statistics – Simulation and Computation*, 38: 6, 1228–1234. <https://doi.org/10.1080/03610910902859574>
- Bandura, A. (1997) *Self-Efficacy: The Exercise of Control*. New York: Macmillan.
- Bellier, H., Bokobza, L., Forsyth, E., Harjani, T. and Taylor, P. (2024) 'KS1 Maths Teacher Choices / One Step Ahead: Evaluation Study Plan'. London: Education Endowment Foundation. Available at: [https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/OSA\\_Study\\_Plan\\_BIT\\_Dec\\_2024.pdf?v=1754982773](https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/OSA_Study_Plan_BIT_Dec_2024.pdf?v=1754982773) (accessed 15 July 2025).
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J. and Balain, S. (2007) 'A Conceptual Framework for Implementation Fidelity'. *Implementation Science*, 2: 40, 1–9. <https://doi.org/10.1186/1748-5908-2-40>
- Data Protection Act 2018, c.12. Available at: [www.legislation.gov.uk/ukpga/2018/12/contents](http://www.legislation.gov.uk/ukpga/2018/12/contents) (accessed 10 June 2024).
- De Corte, E., Mason, L., Depaepe, F. and Verschaffel, L. (2011) 'Self-Regulation of Mathematical Knowledge and Skills'. In: Zimmerman, B.J. and Schunk, D.H. (eds.) *Handbook of Self-Regulation of Learning and Performance*. New York: Routledge, pp. 155–172.
- Department for Education (DfE). (2025a) 'Areas of Research Interest'. London: Department for Education. Available at: [https://assets.publishing.service.gov.uk/media/68136b1bb0ef2c985052541c/DfE\\_areas\\_of\\_research\\_interest.pdf](https://assets.publishing.service.gov.uk/media/68136b1bb0ef2c985052541c/DfE_areas_of_research_interest.pdf) (accessed 18 June 2025).
- Department for Education (DfE). (2025b) 'Schools, Pupils and Their Characteristics, Academic Year 2024/25'. GOV.UK. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics/2024-25> (accessed 05 August 2025).
- Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L.S., Feinstein, L., Engel, M., Brooks-Gunn, J. and Sexton, H. (2007) 'School Readiness and Later Achievement'. *Developmental Psychology*, 43: 6, 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.1428>
- Education Endowment Foundation (EEF). (2020) 'Improving Mathematics in the Early Years and Key Stage 1: Guidance Report'. London: Education Endowment Foundation. Available at: [https://d2tic4wvo1iusb.cloudfront.net/production/eef-guidance-reports/early-maths/EEF\\_Maths\\_EY\\_KS1\\_Guidance\\_Report.pdf?v=1716900941](https://d2tic4wvo1iusb.cloudfront.net/production/eef-guidance-reports/early-maths/EEF_Maths_EY_KS1_Guidance_Report.pdf?v=1716900941) (accessed 15 July 2025).
- Education Endowment Foundation (EEF). (2021) 'Small Group Tuition'. Education Endowment Foundation. Available at: <https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/small-group-tuition> (accessed 15 July 2025).
- Education Endowment Foundation (EEF). (2022) 'Statistical Analysis Guidance for EEF Evaluations'. London: Education Endowment Foundation. Available at: <https://d2tic4wvo1iusb.cloudfront.net/production/documents/evaluation/evaluation-design/EEF-Analysis-Guidance-Website-Version-2022.14.11.pdf?v=1699621596> (accessed 13 August 2025).
- Education Endowment Foundation (EEF). (2025) 'Deployment of Teaching Assistants Guidance Report'. Available at: [https://d2tic4wvo1iusb.cloudfront.net/production/eef-guidance-reports/teaching-assistants/deployment\\_of\\_teaching\\_assistants\\_-\\_guidance\\_report\\_v1.1.0.pdf](https://d2tic4wvo1iusb.cloudfront.net/production/eef-guidance-reports/teaching-assistants/deployment_of_teaching_assistants_-_guidance_report_v1.1.0.pdf) (accessed 02 September 2025).
- Education Policy Institute (EPI). (2023) 'Annual Report 2023: Disadvantage'. Available at: <https://epi.org.uk/annual-report-2023-disadvantage/> (accessed 13 October 2025)

- General Data Protection Regulation (GDPR). (2016) 'Council Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (United Kingdom General Data Protection Regulation) (Text with EEA relevance)'. Available at: [www.legislation.gov.uk/eur/2016/679](http://www.legislation.gov.uk/eur/2016/679) (accessed 02 June 2024).
- GOV.UK (2025) 'Pupil Characteristics – Number of Pupils by Age and Sex, Academic Year 2024/25'. GOV.UK. Available at: <https://explore-education-statistics.service.gov.uk/data-catalogue/data-set/77d37133-ab0e-4428-aa62-9aae5bfdfd73> (accessed 05 August 2025).
- Hembree, R. (1990) 'The Nature, Effects, and Relief of Mathematics Anxiety'. *Journal for Research in Mathematics Education*, 21: 1, 33–46. <https://doi.org/10.2307/749455>
- Jordan, N.C., Kaplan, D., Ramineni, C. and Locuniak, M.N. (2009) 'Early Math Matters: Kindergarten Number Competence and Later Mathematics Outcomes'. *Developmental Psychology*, 45:3, 850–867. <https://doi.org/10.1037/a0014939>
- Krinzinger, H., Kaufmann, L. and Willmes, K. (2009) 'Math Anxiety and Math Ability in Early Primary School Years'. *Journal of Psychoeducational Assessment*, 27: 3, 206–225. <https://doi.org/10.1177/0734282908330583>
- Kroesbergen, E.H. and Van Luit, J.E.H. (2003) 'Mathematics Interventions for Children with Special Educational Needs: A Meta-Analysis'. *Remedial and Special Education*, 24: 2, 97–114. <https://doi.org/10.1177/07419325030240020501>
- Lalley, J.P. and Miller, R.H. (2006) 'Effects of Pre-Teaching and Re-Teaching on Math Achievement and Academic Self-Concept of Students With Low Achievement in Math'. *Education*, 126:4, 747–755. <https://eric.ed.gov/?id=EJ765791>
- Nunes, T., Bryant, P., Evans, D. and Barros, R. (2015) 'Assessing Quantitative Reasoning in Young Children'. *Mathematical Thinking and Learning*, 17: 2–3, 178–196. <https://doi.org/10.1080/10986065.2015.1016815>
- Nunes, T., Barros, R. and Strand, S. (2017) 'Statistical Analysis Plan for 1stClass@Number'. London: Educational Endowment Foundation. Available at: [https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Round\\_8\\_-\\_1st\\_class\\_at\\_number\\_SAP.pdf?v=1716289977](https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Round_8_-_1st_class_at_number_SAP.pdf?v=1716289977) (accessed: 24 July 2025).
- Petronzi, D., Staples, P., Sheffield, D. and Hunt, T.E. (2019) 'Further Development of the Children's Mathematics Anxiety Scale UK (CMAS-UK) for Ages 4–7 Years'. *Educational Studies in Mathematics*, 100: 3, 231–249. <https://doi.org/10.1007/s10649-018-9860-1>
- Ritchie, J. and Spencer, L. (1994) 'Qualitative data analysis for applied policy research'. In: Bryman, A. and Burgess, R. (eds.) *Analysing Qualitative Data*. London: Routledge, pp. 173–194.
- Sahasranaman, A., Morton, C., Bradley, E., Moore, E. and Lord, P. (2025) '1stClass@Number 1: Effectiveness trial: Evaluation Report'. Available at: [https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/1stClassNo.1\\_evaluation-report\\_v.1.0.0\\_2025-06-25-173138\\_pttn.pdf?v=1757322326](https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/1stClassNo.1_evaluation-report_v.1.0.0_2025-06-25-173138_pttn.pdf?v=1757322326) (accessed 02 September 2025).
- Slavin, R.E., Lake, C. and Groff, C. (2009) 'Effective Programs in Middle and High School Mathematics: A Best-Evidence Synthesis'. *Review of Educational Research*, 79: 2, 839–911. <https://doi.org/10.3102/0034654308330968>
- Trundle, R. (2018) 'Changing Lives and Providing Equity Through Pre-Teaching and Assigning Competence'. Leicestershire: Association of Teachers of Mathematics. Available at: <https://atm.org.uk/write/MediaUploads/Journals/MT262/MT26213.pdf> (accessed 21 November 2025).
- van den Berg, M., Bosker, R.J. and Suhre, C.J.M. (2017) 'Testing the Effectiveness of Classroom Formative Assessment in Dutch Primary Mathematics Education'. *School Effectiveness and School Improvement*, 29: 3, 339–361. <https://doi.org/10.1080/09243453.2017.1406376>

- Vaughn, S., Linan-Thompson, S. and Hickman, P. (2003) 'Response to Instruction as a Means of Identifying Students with Reading/Learning Disabilities'. *Exceptional Children*, 69: 4, 391–409.
- Wanzek, J. and Vaughn, S. (2007) 'Research-Based Implications From Extensive Early Reading Interventions'. *School Psychology Review*, 36: 4, 541–561. <https://doi.org/10.1080/02796015.2007.12087917>
- Watt, S.J. and Therrien, W.J. (2016) 'Examining a Preteaching Framework to Improve Fraction Computation Outcomes Among Struggling Learners'. *Preventing School Failure: Alternative Education for Children and Youth*, 60: 4, 311–319. <https://doi.org/10.1080/1045988X.2016.1147011>

## Appendix A: Security classification of trial findings

### OUTCOME: Quantitative Reasoning Test (mathematical reasoning skills)

Rating	Criteria for rating	MDES	Attrition	Initial score		Adjust		Final score
	<b>Design</b>							
5	Randomised design	<= 0.2	0-10%					
4	Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)	0.21 - 0.29	11-20%	4				
3	Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism)	0.30 - 0.39	21-30%			Adjustment for threats to internal validity [-1]		3
2	Design for comparison that considers selection only on some relevant confounders	0.40 - 0.49	31-40%					
1	Design for comparison that does not consider selection on any relevant confounders	0.50 - 0.59	41-50%					
0	No comparator	>=0.6	>50%					

Threats to validity	Risk rating	Comments
<b>Threat 1: Confounding</b>	Moderate (bias could dilute effect)	Randomisation was appropriate and conducted by an independent statistician. Imbalance was moderate in the pre-test (0.06), but it was controlled for in the regression model. All other characteristics were fairly balanced between groups with the exception of gender. Additional analysis included all covariates founds similar results.
<b>Threat 2: Concurrent interventions</b>	Low	Teacher survey highlighted low risk of concurrent interventions and design allowed for existing practices to continue.
<b>Threat 3: Experimental effects</b>	Moderate (bias could dilute effect)	Up to 24% reporting pre-teaching in comparison group from one data source, so potential for spillover is a moderate risk. Trial design with pupil level randomisation increased the risk of experimental effects.
<b>Threat 4: Implementation fidelity</b>	Moderate (bias could dilute effect)	72% reported challenges with implementation and evaluators note inconsistent delivery. This affected the frequency, timing and whether the focus was on pre-teaching.
<b>Threat 5: Missing data</b>	Low	The levels of missing data (beyond attrition) are relatively low and within the acceptable amount specified in the design.
<b>Threat 6: Measurement of outcomes</b>	Moderate (bias unknown)	Teacher administered, with teachers expressing concerns about suitability. However, the measure was previously tested and it appears to have been administered properly.
<b>Threat 7: Selective reporting</b>	Low	No evidence of selective reporting.

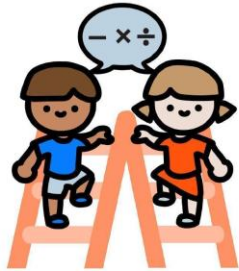
- **Initial padlock score:** 4 Padlocks – This was a well conducted randomised trial With MDES at randomisation of 0.125. However, there was 16% attrition from the trial.
- **Reason for adjustment for threats to validity:** [-1] Padlocks – There are four moderate effects, three of which may reduce the effect size, one of which the direction of impact is unknown.
- **Final padlock score:** initial score adjusted for threats to validity = 3 Padlocks.

## Appendix B: Effect size estimation

Appendix B Table 1: Effect size estimation

Outcome	Unadjusted differences in means	Adjusted differences in means	Pre-teaching group		Non-pre-teaching group		Pooled variance
			n (missing)	Variance of outcome	n (missing)	Variance of outcome	
QRT endline score	-0.019	0.009	1,079 (198)	1.04	1,107 (198)	0.96	1.00
CMAS-UK	0.005	0.004	1,050 (255)	0.124	1,080 (224)	0.120	0.122

## Appendix C: Teacher Guide

 <p>Education Endowment Foundation BIRMINGHAM INSIGHTS TEAM</p> <h1>One Step Ahead Teacher Guide</h1>	<h2>Content</h2> <table border="1"> <tr><td>Introduction</td><td>3</td></tr> <tr><td>Logistics &amp; Set-up</td><td>4</td></tr> <tr><td>Checklist: Planning and Delivering Pre-Teaching Sessions</td><td>5</td></tr> <tr><td>Tips and Guidance</td><td>7</td></tr> <tr><td>Dos and Don'ts</td><td>9</td></tr> <tr><td>Examples of Pre-Teaching Activities &amp; Strategies</td><td>10</td></tr> <tr><td>Storytelling Example</td><td>13</td></tr> <tr><td>FAQs</td><td>14</td></tr> <tr><td>Resources</td><td>17</td></tr> </table>	Introduction	3	Logistics & Set-up	4	Checklist: Planning and Delivering Pre-Teaching Sessions	5	Tips and Guidance	7	Dos and Don'ts	9	Examples of Pre-Teaching Activities & Strategies	10	Storytelling Example	13	FAQs	14	Resources	17
Introduction	3																		
Logistics & Set-up	4																		
Checklist: Planning and Delivering Pre-Teaching Sessions	5																		
Tips and Guidance	7																		
Dos and Don'ts	9																		
Examples of Pre-Teaching Activities & Strategies	10																		
Storytelling Example	13																		
FAQs	14																		
Resources	17																		

<h2>Introduction</h2> <p><b>Summary</b></p> <p>One of the best ways to enable meaningful learning is to ensure that all pupils are able to access the key content of each lesson, and approach the lesson with a positive mindset and confidence in their abilities. Teachers have many ways to do this, including through targeted out-of-class support. This can happen before or after the main lesson. There is no evidence-based consensus over which is more effective.</p> <p>The One Step Ahead study aims to test the effectiveness of providing regular support for small groups of pupils before the main maths lesson in Key Stage 1.</p> <p><b>How this works:</b> You'll deliver short (around 15mins) pre-teaching sessions in maths;</p> <ul style="list-style-type: none"> <li>Before the main lesson;</li> <li>At least twice a week;</li> <li>To pupils in the 'pre-teaching group', using the list of pupils we will send you;</li> <li>For one term (September to December 2024). Consistency is key for impact.</li> </ul> <p><b>Why this works:</b> Pre-teaching is known to help pupils make connections between what they are learning and what they already know. By warming up prior knowledge, through previewing lesson content or building cognitive scaffolds, we allow pupils to approach the content of the main lesson more confidently, improving their mathematical understanding and their sense of self-efficacy. It also enables them to go into the lesson with a more confident and positive mindset, by introducing maths as something they can do. This can also have positive effects on the rest of the class.</p> <p><b>How to Use this Guide?</b></p> <ul style="list-style-type: none"> <li>As a reference to understand the aims of the study: Information about the objectives of the study, and why and how we think it will have a positive impact on pupils.</li> <li>As a summary of instructions, including what to do (and not to do) during the pre-teaching.</li> <li>As a tool to help you plan your pre-teaching sessions: You'll find a checklist to help you prepare for pre-teaching, as well as tips and guidance to support the implementation.</li> <li>As a source of inspiration: We're providing a (non-exhaustive) list of activities you can do during the pre-teaching.</li> <li>As a resource you can come back to during the study, to get new ideas for pre-teaching sessions, and to help you reflect on pupils' progress.</li> <li>As a first port call if you have any questions on the project or sessions. If you have any questions, feel free to get in touch by emailing <a href="mailto:1stepahead@bi.team">1stepahead@bi.team</a>.</li> </ul> <p>This guide is not a scheme of work or a set of lesson plans. It's up to you to decide what to do in your pre-teaching slots. Whatever you choose to do should be simple and should not require a large amount of planning.</p>	<h2>Logistics &amp; Set-up</h2> <p><b>Who receives pre-teaching?</b></p> <ul style="list-style-type: none"> <li>Before the summer holidays, we asked you to identify a group of 12 pupils in your class that could benefit from additional support and more active engagement in maths.             <ul style="list-style-type: none"> <li>We have randomly assigned half of these pupils to be in the 'pre-teaching group', who will be receiving the pre-teaching sessions this term.</li> <li>The pre-teaching group should remain fixed throughout the study period (those same 6 pupils should be getting the pre-teaching sessions each week).</li> </ul> </li> <li>Don't do anything differently with other pupils in the class, including the other six pupils you had identified as eligible for One Step Ahead. The latter are the comparison group for the study, and they are key to measuring impact.</li> <li>Other pupils should not receive any maths pre-teaching, and otherwise continue to benefit from the in-class and out-of-class support they would usually be getting.</li> </ul> <p><b>Who delivers pre-teaching?</b></p> <ul style="list-style-type: none"> <li>The sessions should ideally be delivered by the class teacher (or the person teaching the main class).             <ul style="list-style-type: none"> <li>Where the teacher is not able to deliver the sessions, ideally the person doing so (e.g. the Teaching Assistant) should also be present in the main lesson, and should take notes during the pre-teaching session for the teacher.</li> </ul> </li> </ul> <p><b>What does it consist of?</b></p> <ul style="list-style-type: none"> <li>10-15 mins pre-teaching sessions focusing on the content of the upcoming lesson(s) in maths.</li> <li>The content and structure of the sessions is up to you, but you can find some inspiration and ideas in this guide.</li> </ul> <p><b>When does it take place?</b></p> <ul style="list-style-type: none"> <li>At least twice a week, for one term.</li> <li>Before the main lesson.             <ul style="list-style-type: none"> <li>It should be delivered as close to the main maths lesson as possible, ideally on the day of the main lesson where feasible.</li> </ul> </li> </ul>
--	--



## Checklist: Planning and Delivering Pre-Teaching Sessions

### Steps

- 1 **As you plan for the upcoming maths lesson or unit, consider the key concepts that will be developed in the next lesson. Try to pick one.**
- 2 **Diagnosis of need:** What is the main thing that would enable pupils to access the upcoming lesson or unit?
- 3 **Plan your pre-teaching session:** choose one activity **closely aligned with the key concept you've identified.**
- 4 **Deliver the pre-teaching session.**
  - a. Use the **same language** you will use in the main lesson (e.g. sentence stems)
  - b. Keep the activities **close to the key concept** of the main lesson.
- 5 **During the main lesson, try to make active links between the pre-teaching and main lesson:**
  - a. **Use the same language** as you did in the pre-teaching session.
  - b. Encourage pupils to **repeat** comments they made or things they did particularly well during the pre-teaching sessions.
  - c. Use **public praise and encouragement** to foster active engagement.
- 6 **Reflect** on the pre-teaching and main lesson, considering pupils' progress and changes in confidence, attitudes, engagement. Record these notes in writing if you can.

5



## Checklist: Planning and Delivering Pre-Teaching Sessions

### What does a typical pre-teaching session look like?

#### Example: Identifying unit fractions of a quantity

1. **Identify core concepts** for the upcoming lesson:
  - a. To understand and use key vocabulary relating to fractions.
  - b. To find  $\frac{1}{3}$  of an amount.
2. **Diagnosis of need:** What is the main thing that would enable pupils to access the upcoming lesson or unit?
  - a. Familiarity with vocabulary; being able to point out the different parts of a fraction.
  - b. Understand that the whole needs to be split into three equal parts.
3. **Plan key questions and activities:**
  - a. How can I frame the session in an engaging way?
  - b. How can I introduce key vocab such as numerator and denominator? How can I help the pupils to practice this language?
  - c. **Suggested activity:** Using manipulatives.
4. **Delivery (15mins in total):**
  - a. Tell the pupils a short story about a pirate finding treasure 'gems', and model her sharing it with her crew (using physical manipulative representations for each unit).
  - b. Emphasise the key vocab using the manipulatives. Be specific to the language that you're going to use in the main lesson. Consider using sentence stems and choral repetition. E.g. "There are x gems, we divide the gems into three equal parts. I have one of them, we call this a third (see NCETM guidance for KS1)."
  - c. Ask the pupil "pirates" (in pairs / "crews") to share their own gems according to the fractions you show them. Key question: "What is a third of 6 gems? How did you work this out?"
5. **During the main lesson:** make **active links** with the pre-teaching
  - a. Example of **active link** between pre-teaching and main lesson: "[Pupil A] was great at figuring out thirds this morning. Do you remember how we were counting gems? Can you explain to the person next to you how you did this?"
6. **Reflect:**
  - a. How did the pupils respond to the activities (e.g. manipulatives)?
  - b. Is there a type of session structure / activity that they engage with particularly well? Is there anything they struggle to grasp, or doesn't keep them engaged as long?
  - c. Did the session affect their behaviour and attitude in the main lesson?
  - d. Did the session prepare them adequately for the main lesson? Did they demonstrate a grasp of the key concepts?

6



## Tips and Guidance

### TOP TIPS

- 1 **Keep it simple!**
- 2 Whatever the activity, **consider language** and **use the same vocabulary** in the pre-teaching and the main lesson.
- 3 In the first pre-teaching session, ask pupils to help you **come up with a fun name for the group**, e.g. 'Sneaky Peek' group, 'Maths Masterclass', 'Brain Boosters'... This helps foster positive perceptions of and engagement with pre-teaching and maths more generally!
- 4 Make **active links** between the pre-teaching and the main lesson.

#### Planning (checklist steps 1, 2 & 3): Some questions you can think of to help you plan...

##### Identifying the key concept:

- What is the most important idea / concept of the upcoming lesson?
- What's the best 'starting point' that will enable children to access this concept?
- What are the best ways to support pupils' understanding of this concept?

##### Enabling children to access the key concepts:

- Which manipulatives or representations will support pupils' understanding of the concept?
- What are the words or key vocabulary you expect pupils to use?
- What activity are you planning pupils to engage in during the session? We recommend one activity per pre-teaching session.

##### Identifying the focus of pre-teaching:

- What are likely common misconceptions? What might pupils get stuck with?
- How can you help pupils avoid or overcome these?

##### Adapting Teaching and Learning:

- How would levels of ability for pupils make your approach different? (see EEF resources on Adaptive Learning)

7



## Tips and Guidance

#### Delivery (checklist step 4)

- **If possible, take notes of contributions/comments** pupils make in pre-teaching sessions. This will help with tracking their progress throughout the study, and making 'active links' with the main lesson.
- If anyone other than the class teacher is delivering the pre-teaching session, make sure they also take notes (focusing on what worked well/not so well during the session, and any changes in pupils' attitudes and behaviours).
- **Frame pre-teaching as something exciting.** Remind pupils that the group is learning things or doing activities that the other pupils have not done yet.

Some teachers find it useful to deliver the sessions following a **regular structure**. This is just a suggestion and may not work for your group, or for every session! Example structure:

1. **Outline** the session's objective, and **demonstrate** the skill or concept you will be covering.
2. **Practice** the skill/vocabulary together with an activity.
3. **Assess** learning and understanding ahead of the main lesson (e.g. with a mini plenary).

#### Outside the pre-teaching session (checklist steps 5 & 6)

##### Here are some tips to make 'active links' between the pre-teaching and the main lesson:

- Use the **same language** in the pre-teaching session and the main lesson.
- Give pupils **ownership of their ideas and comments:**
  - If they're comfortable doing it themselves, ask if they'd like to tell the rest of the class later
  - If they're not comfortable, you can still emphasise and praise their contribution. E.g. "[Pupil name] said something really interesting about this in our Sneaky peek...".
- If pupils are paired up during the lesson, try to pair pupils in the pre-teaching group with higher-attaining pupils. **Reverse the usual roles** of who is helping whom.
- **Celebrate the thinking and the ideas**, rather than the child. E.g. instead of "Paloma was really good at addition this morning", consider "Paloma had an interesting way of thinking about her calculations this morning that I wanted to share with the rest of you".

8

## Dos and Don'ts

### Preparation

Dos	Don'ts
Keep the planning and delivery of the sessions as simple as possible.	Overcomplicate the content of the pre-teaching.
Focus on the key elements you want the children to take away from the session in a short amount of time.	Use this as catch-up / to go over content from the previous lesson.
Ask the pupils to help you come up with a fun name for their pre-teaching group! E.g. 'Sneaky Peaks group', 'Maths Masterclass'...	Try to cover too much in a single session, or teach all the content from the main lesson. Focus on a single aspect of the upcoming lesson, or a single skill.
Actively link the session content to upcoming lesson content e.g. by using the same vocabulary/stems and the same/similar questions and tasks.	
Use visual aids & manipulatives where feasible.	
Use activities that are simple and familiar to pupils to maximise time spent practising rather than 'explaining'.	
Prioritise tasks and activities which have previously engaged pupils' interest.	
Identify key questions that will allow children to focus on the key concepts of the lesson.	

### Implementation

Dos	Don'ts
Frame sessions as an exciting preview of the main lesson.	Frame or structure the sessions in a way that might resemble an 'extra lesson'.
Make it clear that this group now knows things that the rest of the class haven't yet covered; make the sessions feel exciting and special.	Dwell too much on things that the group struggled with during the last lesson.
Make notes of pupils' comments in the sessions that you might want to bring up in the main lesson.	Let the focus stray too much from upcoming content and prerequisite knowledge.
Highlight things they did well in the main lesson.	Put pupils on the spot by asking them to perform something during the main lesson if they are not comfortable with it.
Be generous with positive praise and confidence boosting.	Repeat the pre-teaching session if some pupils missed it.
Ask pupils if they'd be willing to share their contributions from the pre-teaching in the main lesson (or if they want you to share them).	

9

## Examples of Pre-Teaching Activities & Strategies

### Interactive Games related to the upcoming lesson (See Grace Coker - [EEF guidance on using games to promote purposeful practice](#))



Consider using matching and identifying games, for example when teaching geometry / shapes.



Try simple board games which are easy to produce and can be adapted to suit different purposes.



Explore interactive online games - see [NHTS](#) for ideas on how to utilise other resources.

### Use virtual representations and / or manipulatives to present a mathematical idea



Use counters or drawings to visually demonstrate a process: for example, if teaching subtraction, use a physical representation of taking away.



Encourage children to create their own manipulatives or visual representations, and challenge them to explain how this reflects the mathematical idea.



Using manipulatives, ask pupils what they can see or notice, what stays the same, and what changes. E.g. using a blank cherry / part-whole model and counters, ask pupils to add or subtract numbers, explaining what they can see, and what changes when the counters move.

### Addressing misconceptions / Unpicking worked example



In pairs, ask pupils to talk with each other through each step of an example.



Model how you would approach an example, including potential misconceptions or areas you might get stuck, and how you can avoid these.



Ask the group to 'act as the teacher', and instruct you how to go about an example.

10

## Examples of Pre-Teaching Activities & Strategies

### Storytelling



Start with a story that incorporates the mathematical concept that you are about to teach.



You can invent a story based on the concept, or you can use and adapt published texts. See an example story on the next page that can easily be adapted for different concepts.



Encourage pupils to come up with their own story that uses the concepts.

### Predict and discuss: Ask children to predict what they think the upcoming lesson might be about.



Use prompts - images, manipulatives, music, etc. - to stimulate thinking.



Allow pupils to discuss their ideas and explain their reasoning.



Engage in a discussion with them about this topic. What are their expectations? What do they feel confident/less confident about in relation to this lesson?

### Teaching a scaffold / 'Getting unstuck': Introduce a key strategy that will support children to work independently during the lesson.



You can show children how to use concrete manipulatives to help them with an independent task in the main lesson.



Help children to write a bank of number facts to which they can refer independently during the lesson.



Children can also practise using a scaffold they've just learned on questions or concepts similar to those they will encounter during the lesson.

11

## Examples of Pre-Teaching Activities & Strategies

### Plan and teach



Start by telling pupils your lesson plan for the day, and focus on a specific concept you are planning to teach.



Ask the group to help you come up with an activity / game you can do with the whole class to introduce the concept. E.g. how can we explain to the class what 'numerator' are? Can you come up with a fun rhyme we can teach the class later to help them memorise the word?



In the main lesson, ask pupils in the pre-teaching group to post up with other pupils and show their peers the activity (e.g. the rhyme or game they came up with). Try to reverse the roles of which pupils are helping others by pairing the pre-teaching pupils with higher-attaining pupils.

### Making comparisons



Show the pupils different examples of the concept you will be teaching in the main lesson (e.g. images of different shapes, different fractions, sums, etc.) If available, use manipulatives to support the demonstration.



Use phrases such as "Show me... Tell me" to engage pupils:  
 • "Show me an example of a (triangle, denominator, etc.)"  
 • "Can you tell me why this is a triangle?"



Ask them to point out why the examples are different or the same (e.g. One fraction is bigger than the other, one shape is a triangle, the other is a square). Use the sentence stems and vocabulary you will use in the main lesson, e.g. "This is a triangle because it has three sides".

### Boosting mathematical literacy



Try reading some of the mathematical stories (see storytelling example) with the pupils.



Ask the pupils to identify key words and vocabulary in a text. Help them sound them out, and ask them to practise writing them independently. To help them remember spelling, look for words within words e.g. 'rectangle, triangle'



If pupils are struggling to read or write specific vocabulary, help them create a "word bank". This can be a page of their lesson book where they write key words with a mnemonic to help them remember what it means, e.g. a drawing, or a short definition they can come back to independently.

12



## Storytelling example

Below is a template of a simple story you could adapt for the 'Storytelling' activity, with some suggestions on how to use it for different topics. Feel free to make changes based on the topic you are teaching and the needs of pupils in your pre-teaching group.

A group of pirates are sailing the seas trying to find a treasure. They land on a desert island which has so much treasure they can hardly count it! The pirates decide to split off and find their own treasure.

### Unit: Numbers and place value

When they get back to the boat, they've collected so much treasure that they can't count it all in one go.

Maise finds two hands + 2 fingers of gold coins (12 gold coins).

Lou finds 4 hands of gold coins (20 gold coins).

Mo finds 1 hand of gold coins (5 gold coins).

*How many gold coins did Maise / Lou / Mo find? Count on your fingers with a friend.*

When they are counting all their treasure, the pirates put all their coins together. They have 37 coins in total. However, they then decide to split it up again into piles of 10, so they can share it with all their crew later on. *How many tens do they have in total? How many piles of 10 can they split their coins into? How many ones will they have left over?*

### Unit: Addition and subtraction

Maise finds 3 gold coins.

Lou finds 2 gold coins.

Mo finds 3 gold coins.

*How many gold coins do they have in total?*

*Lou accidentally drops 3 gold coins in the sea. How many gold coins do they have left?*

### Unit: Division and fractions

Maise finds 3 gold coins.

Lou finds 5 gold coins.

Mo finds 4 gold coins.

*How many gold coins do they have in total?*

*The pirates want to share their treasure equally among themselves. How can we represent this with fractions?*

*How many coins do they get each?*

13

## FAQs: Planning & logistics

**When should I schedule the sessions?**

**Pre-teaching sessions should take place on the morning of, or on the day before, the main maths lesson.**

Beyond that, it's completely up to you when in the week or at what time in the day the sessions happen, so feel free to schedule them whenever works best for you, so long as they take place before the main maths lesson that the pre-teaching is covering. Some teachers find the best time to do the pre-teaching is during morning assembly or form time.

**How much time should I spend planning a session?**

**Planning for each pre-teaching session should not take very long (10 to 15 minutes at most)**

We recommend incorporating this in your normal lesson planning practice; see the Planning and Delivery Checklist for some tips on how to do this.

**What if we have to cancel a session due to unforeseen circumstances?**

**Please reschedule any one-off missed sessions if this is feasible in the school timetable, to ensure pupils receive the minimum of two pre-teaching sessions per week.**

We know schools are busy places and things don't always go according to plan! If sessions are missed on a more regular basis, please do let us know by emailing us at [1stepahead@bi.team](mailto:1stepahead@bi.team) and in the short surveys we send you throughout the study.

**How are other schools arranging cover for the rest of the class during the pre-teaching session?**

Most schools that took part in the pilot for this project **scheduled the sessions in the morning, during form time, assembly, or reading time.** Each school may have a different approach based on the timings of their school day, and availability of cover.

**We do more than two maths lessons a week. Can we pick before which the pre-teaching takes place?**

**Yes.** As long as it happens before the lesson you are pre-teaching, you can choose which days and frequency works best.

14

## FAQs: Pre-teaching group composition

**What should I do if a pupil in the pre-teaching group is not showing signs of progress during the study?**

**This can happen!** Pupils may progress in different ways and at a different pace, and you may not always notice an immediate boost in their confidence or understanding.

**If this happens, please do not remove any pupils from the pre-teaching group** (unless they are showing signs that taking part in the sessions is causing them harm or distress). For pre-teaching to have an impact in the medium and longer term, pupils should benefit from regular sessions for at least one whole term.

**What should I do if a pupil in the pre-teaching group has fully understood the concept I was planning on pre-teaching to the others, or does not seem to need the pre-teaching anymore?**

**The pre-teaching group needs to stay fixed throughout the study** in order for us to accurately measure the impact of pre-teaching. This will also maximise impact.

This means that even if a child is showing signs of having understood the content you are covering in the sessions, it's important that you **do not remove them from the group for the duration of the study.**

**Can I add pupils to the pre-teaching group if I think they would also benefit from the content?**

**No, the pre-teaching group needs to stay fixed throughout the study** in order for us to be able to accurately measure the impact of pre-teaching. This will also maximise impact.

Once pupils have completed the assessments at the end of the study in December, you can change the group as you see fit.

15

## FAQs: Session structure and content

**I already run catch-up sessions for maths. Should the content I cover in the pre-teaching be different?**

**Pre-teaching should focus on what they are likely to struggle with in the upcoming lesson(s), rather than on what they have struggled with in previous ones.**

There may be some crossover between the two when pre-teaching covers topics that span across more than one lesson, but pre-teaching should always be forward-looking.

**Can I use the same pre-teaching activity more than once?**

**Yes!** We encourage you to reuse activities and strategies in different pre-teaching sessions, especially if the pupils reacted well to them.

16

# Appendix D: Information and withdrawal form for caregivers



## Research Project Privacy Notice for school staff and parents

One Step Ahead: A Teacher Choices project

### Introduction

Your school or your child's school is participating in a research project testing the impact of an early learning Teacher Choices project called **One Step Ahead**. A common question teachers ask about their practice is how to best support pupils to access lesson content in mathematics. One Step Ahead explores how regular targeted pre-teaching sessions for Key Stage 1 pupils can improve their confidence and mathematical reasoning skills, compared with other usual approaches to targeted support that do not involve pre-teaching.

The project is a collaboration between the Behavioural Insights Team (BIT) and the Education Endowment Foundation (EEF). BIT has designed the project, and will be delivering and evaluating the project. The EEF are the funders of the project, including its evaluation. In each participating school, 24 pupils will take part in the initiative (12 in Year 1 and 12 in Year 2). In each year-group, 6 of the 12 pupils will receive regular pre-teaching maths sessions over one term, and the other 6 will continue to benefit from the in-class and out-of-class support they would usually receive, but not receive any pre-teaching.

This privacy notice sets out how BIT, as a data controller, will collect and use the personal data of the pupils and school staff that participate in the research.

### Contact details

Behavioural Insights Ltd (the legal name of BIT is the controller and is responsible for any personal data collected in connection with this evaluation. We have appointed a Data Protection Officer (DPO) who is responsible for overseeing questions in relation to this privacy notice. If you have any questions about this privacy notice, including any requests to exercise a school staff member or a pupil's legal rights in relation to staff or a pupil's personal data, please contact the DPO:

**Post:** Behavioural Insights Ltd, 58 Victoria Embankment, London EC4Y 0DS.

**Email:** dpo@bit.team.

1

You also have the right to make a complaint at any time to the Information Commissioner's Office (ICO), the UK supervisory authority for data protection issues ([www.ico.org.uk](http://www.ico.org.uk)). We would, however, appreciate the chance to deal with your concerns before you approach the ICO so please contact us in the first instance.

## What personal data will we collect and what will we do with it?

Most personal data will be collected directly from participating schools. Data will be shared and stored securely, in accordance with the Data Protection Act (2018) and UK GDPR (General Data Protection Regulation).

For participating Year 1 and Year 2 pupils, the following data will be collected:

- Full name
- Date of birth
- Gender
- Free School Meals eligibility status
- English as an Additional Language (EAL) status
- SEN status
- Type of SEN (if applicable)
- Unique Pupil Number (UPN)
- Year group
- Attendance to pre-teaching maths sessions
- Results from a short mathematical assessment at the start and end of the trial
- Results from a short questionnaire about maths anxiety at the end of the trial
- Focus group discussion data - qualitative data gathered in focus groups on experiences of the pre-teaching sessions\*
- Observation fieldwork notes\*
- School name and URN

\* This data will only be collected for pupils in schools selected for case studies and observations and will therefore not be collected for every pupil involved in the One Step Ahead project. You will be informed if your school or your child is selected for focus groups and observations, which are voluntary opt-in activities for schools.

BIT will use the data collected to conduct an evaluation of the One Step Ahead trial. As part of this evaluation, we are running a trial in approximately 95 primary schools in England. This trial will involve Year 1 and Year 2 teachers in all participating schools to identify 12 pupils in their class who may be eligible for pre-teaching sessions in maths. BIT will randomly allocate half of the identified pupils to receive pre-teaching sessions ('pre-teaching group'). The other half of the identified pupils will not receive pre-teaching sessions in KS1 Maths, but can receive any other type of additional support typically provided (forming a 'comparison' group).

At the end of the project, BIT will use the pupil data that it has collected (as listed above) to measure the effect of One Step Ahead pre-teaching on pupils' skills and confidence in maths. BIT will also assess how the impact (if any) is achieved.

2

For school staff (including the key contact person, Year 1 and Year 2 teachers, and Maths lead where relevant), the following data will be collected:

- Name
- Email address
- Job title
- Number of years' teaching experience
- Specialism (if any)
- Usual practice of providing targeted or additional out-of-class support for KS1 Maths
- Lesson topics covered in the term
- Class size
- Attitudes and views about the pre-teaching approach and the trial
- Attitudes and views about the guidance materials for the trial

This data will be used to deliver the initiative. Specifically, BIT will use school staff data to:

- Communicate directly with the school staff to facilitate delivery of the One Step Ahead initiative.
- Share this information with the project funders, EEF, who will use it to monitor the recruitment of schools to the project.
- Enable BIT to recontact school staff to invite them to take part in future research projects.

We will not use any identifiable staff, pupil, or school data and information in any report that we publish about the One Step Ahead project.

### Withdrawing pupil participation

At any point until July 15, 2024, you can withdraw your child's participation from the trial. If you withdraw your child from the trial before any of their personal data is shared with BIT, then schools will not share your child's personal data, your child will not participate in the pre-teaching sessions, and they will not take the baseline and endline tests. Should we be informed that you would like to withdraw your child from the trial after some personal data has been collected, we will not collect any additional data for your child. Wherever possible, we will delete all of your child's personal data collected up to that point (other than interview responses that have already been anonymised and analysed, which will be handled in accordance with the remainder of this privacy notice).

### What is our lawful basis for processing your personal data?

Data protection laws require us to meet certain conditions before we are allowed to use staff and pupil data in the manner described in this notice, including having a lawful basis for the processing.

3

**For all information collected as part of the One Step Ahead project, BIT is relying on the lawful basis of:**

**LEGITIMATE INTERESTS:** Our lawful basis for processing the personal data listed above is legitimate interests (as per Article 6 (1) (f) of the GDPR) and we have considered that staff and pupil interests and fundamental rights do not override those legitimate interests. It is necessary in BIT's 'legitimate interests' to process the personal data identified above in order to learn more about the teaching of mathematics in the classroom to help us assess the feasibility of different teacher choice approaches that aim at improving maths attainment and reducing attainment gaps between pupils. The research project fulfils BIT's core business aims including undertaking research, evaluation and information activities in sectors that will deliver social impact. This determination is based on an assessment undertaken by BIT's Legal team, which is performed in advance of each new research project.

### Who has access to your information?

Staff and pupil information will be accessed by a limited number of researchers and advisors in BIT's project team working on this project.

BIT may disclose the information to third parties in connection with the purposes of processing the personal data set out in this notice. These third parties may include:

- Other companies in BIT's group (that are based in the United Kingdom)
- Regulators, law enforcement bodies and the courts, in order to comply with applicable laws and regulations, assist with regulatory enquiries, and cooperate with court mandated processes, including the conduct of litigation;
- Suppliers, research assistants and sub-contractors who may process information on behalf of BIT e.g. cloud services to store data, transcription services, and survey platforms. These third parties are known as data processors and when we use them we have contractual terms and policies and procedures in place to ensure that personal data is protected. This does not always mean that they will have access to information that will directly identify school staff or pupils as we will share anonymised or pseudonymised data only wherever possible. We remain responsible for staff and pupil personal information as the controller.

Anonymised data may be made available to other researchers and organisations.

We may also disclose staff or pupil personal information if required by law, or to protect or defend ourselves or others against illegal or harmful activities, or as part of a reorganisation or restructuring of our organisations.

For the purpose of research and archiving, pupil-level impact data (e.g. maths learning and anxiety assessment results) and some pupil personal data (e.g. name, date of birth, UPN) that BIT has collected will be securely shared with the EEF's Data Archive manager (FFT). FFT will match each pupil's personal data with a unique Pupil Matching Reference number in the Department for Education's (DfE) National Pupil Database. Once matched, all pupil personal data that could directly identify a pupil will be deleted and the remaining pseudonymised pupil data will be added to the EEF Archive. At this point, EEF will become

4

the controller of the data transferred to the EEF Data Archive. The EEF Data Archive is held in the Office for National Statistics (ONS) Secure Research Service and includes information that is only identifiable to the DfE. The data is kept indefinitely for the purposes of future research, which may entail matching the data with other administrative data. For more information about the EEF Data Archive and EEF's privacy notice, please see [here](#).

### International transfers

Staff and pupil personal information will not be transferred outside of the European Economic Area ("EEA"). References in this notice to the EEA include the UK, even where the UK is no longer a member of the European Union / European Economic Area.

### Security

We take reasonable steps to protect personal information and follow procedures designed to minimise unauthorised access, alteration, loss or disclosure of personal information.

Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, we implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk of processing.

We ensure that those who have permanent or regular access to personal data, or that are involved in the processing of personal data, are trained and informed of their rights and responsibilities when processing personal data. We provide such access on a need-to-know basis, and have measures in place which are designed to remove that access once it is no longer required.

Physical personal devices used by BIT are encrypted to protect personal data, and confidential hard copy data (including special category data) is kept in locked rooms or cabinets.

We have put in place procedures to deal with any suspected personal data breach and will notify staff and parents and any applicable regulator of a breach where we are legally required to do so.

### Data retention

We will only retain staff and pupil personal data for as long as necessary to fulfil the purposes we collected it for, including for the purposes of satisfying any legal, accounting, or reporting requirements. When it is no longer necessary to retain the staff and pupil personal data, it will be securely deleted.

To determine the appropriate retention period for personal data, we consider the amount, nature, and sensitivity of the personal data, the potential risk of harm from unauthorised use or disclosure of the personal data, the purposes for which we process the personal data and whether we can achieve those purposes through other means, and the applicable legal requirements.

5

Taking the above factors into consideration, the approximate date of deletion for staff and pupil personal data is 31st December 2025 (within three months of the final report for the One Step Ahead project being published).

In some circumstances, we will retain an anonymised dataset (so that it can no longer be associated with staff or pupils) for research or statistical purposes, in which case we may use this information indefinitely without further notice.

### School staff and pupil legal rights

Under certain circumstances, staff and pupils have rights under data protection laws in relation to their personal data, including rights to:

- Request access to your personal data: This enables you to receive a copy of the personal data we hold about you and to check we are lawfully processing it.
- Request correction of your personal data: This enables you to have any incomplete or inaccurate data we hold about you corrected.
- Request erasure of your personal data: This enables you to ask us to delete or remove personal data where there is no good reason for us continuing to process it.
- Object to processing of your personal data: For example, you can object where we are relying on a legitimate interest (or those of a third party) and there is something about your particular situation which makes you want to object to processing on this ground as you feel it impacts on your fundamental rights and freedoms.
- Request restriction of processing your personal data: This enables you to ask us to suspend the processing of your personal data.
- Data portability: Where the processing takes place on the basis of your consent or contract, and is carried out by automated means, you have the right to request that we provide your personal data to you in a machine-readable format, or transmit it to a third party data controller, where technically feasible.
- Withdraw consent to the processing of your personal data: This applies where we have relied on consent to process personal data. Please note that withdrawal of consent will not affect the lawfulness of any processing carried out before withdrawing your consent.
- Not be subject to decisions based purely on automated processing where it produces a legal or similarly significant effect on you. Please note that BIT does not engage in automated decision making without manual intervention in its research projects.

If a staff member or a pupil wishes to exercise any of the rights set out above, please contact the Data Protection Officer with your specific request by email to: [dpo@bit.team](mailto:dpo@bit.team)

It is important to understand that the extent to which these rights apply to research will vary and that in some circumstances your rights may be restricted.

Ordinarily, you will not have to pay a fee to access your personal data (or to exercise any of the other rights). However, we may charge a reasonable fee if your request is clearly unfounded, repetitive or excessive. Alternatively, we may refuse to comply with your request in these circumstances.

6

We may need to request specific information from you to help us confirm your identity and ensure your right to access your personal data (or to exercise any of your other rights). This is a security measure to ensure that personal data is not disclosed to any person who has no right to receive it. We may also contact you to ask you for further information in relation to your request to speed up our response.

We try to respond to all legitimate requests within one month. Occasionally it may take us longer than a month if your request is particularly complex or you have made a number of requests. In this case, we will notify you and keep you updated.

Please also note that we can only comply with a request to exercise your rights during the period for which we hold personal information that directly identifies you. If we have only collected pseudonymised information (e.g. where we have not collected any names or contact details) or personal data has been irreversibly anonymised and has become part of the research data set, it will not be possible for us to comply. Please note that this does preclude you exercising your right to complain via the Information Commissioner's Office ([www.ico.org.uk](http://www.ico.org.uk)).

### Changes to this notice

We may change this [Privacy Notice](#) from time to time. If we make any significant changes in the way we treat staff and pupil personal information we will make this clear by updating the notice on the [project website](#) or by contacting staff and parents directly.

7

## Appendix E: Overall guidance for teachers



### Teacher Assessment Guidance for One Step Ahead

#### What?

- There are two assessments:
  - The **Quantitative Reasoning Test (QRT)** measures children's mathematical reasoning skills.
  - The **Children's Mathematical Anxiety Scale (CMAS)** measures children's anxiety related to maths.
- Physical copies of the tests will be sent directly to your school.
- You will administer the tests to pupils. As with the baseline assessments, you can then arrange for DHL couriers to collect from your school, so that they will be marked externally.
- You will receive guidance to help you administer each assessment.
  - This includes instructions for administering the QRT (separate instructions for Year 1 and Year 2), and for administering the CMAS.
  - You will also receive access to an online training video for the QRT, as before with the baseline assessments. If you would like to refresh your memory, log on to the Sharepoint using your email address and the link provided and watch a short (~30-min) training video on how to administer the QRT.

#### Who?

- Please administer both tests to the 12 pupils in each class who have been selected as eligible for the pre-teaching sessions.
  - All pupils initially identified as eligible must receive the assessments, even if they were in the comparison group who did not receive pre-teaching this term.
- You do not need to deliver the tests to the rest of the class.
- For Y1 pupils, the QRT will need to be administered individually, whereas the CMAS can be administered to the whole group.
- For Y2 pupils, both the QRT and CMAS can be administered to the whole group.

#### When?

- In the last three weeks of the Autumn term, follow the handbook instructions and administer the QRT using Form B to the 12 pupils in your class that were nominated for pre-teaching.
  - [Y1 teacher] Administer the test to each pupil individually
  - [Y2 teacher] Administer the test to the group of 12 pupils
- At the end of the Autumn term, administer the CMAS to the group of 12 pupils that were nominated in your class as eligible to receive pre-teaching
- Give the completed test sheets for both tests to your school's project lead

#### Why?

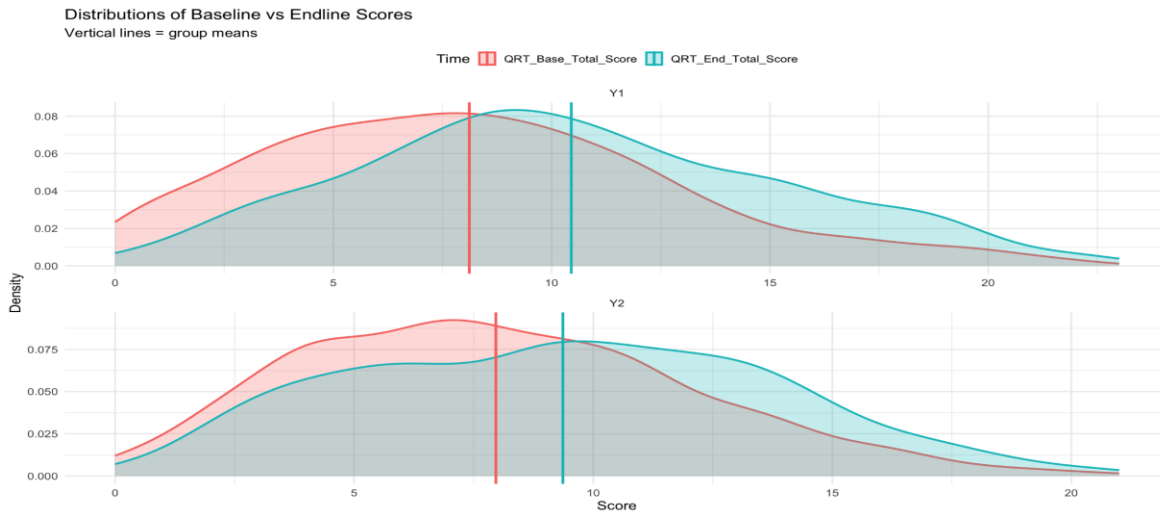
- To measure the effect of pre-teaching on the maths reasoning skills and confidence of your pupils.

**Please follow the instructions as closely as you can.**

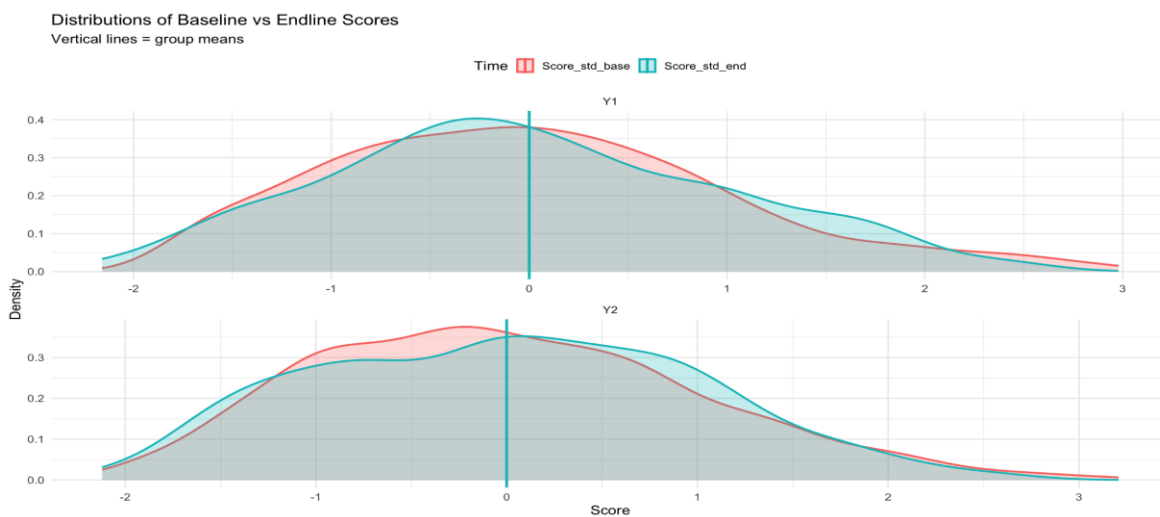
If you have any questions, talk to your school's project lead or email us at [1stepahead@bi.team](mailto:1stepahead@bi.team).

## Appendix F: Distribution of outcome measures

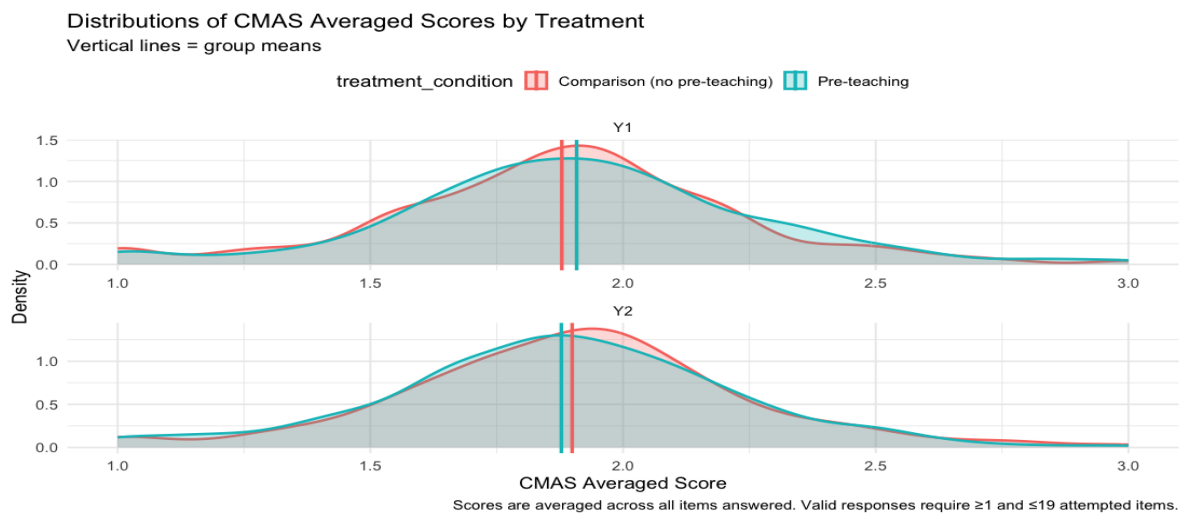
Appendix F Figure 1: Distribution of raw, unstandardised QRT scores (primary outcome) at baseline and endline by treatment group



Appendix F Figure 2: Distribution of standardised QRT scores (primary outcome) at baseline and endline by treatment group



Appendix F Figure 3: Distribution of average CMAS-UK (secondary outcome) at endline by treatment group



## Appendix G: Regression outputs for primary and secondary analysis

Appendix G Table 1: Primary analysis

	Dependent variable: Standardised QRT endline score
Treatment	0.009 (0.034)
Standardised QRT baseline score	0.545*** (0.02)
Constant	-0.085 (0.098)
Observations	2,186
R2	0.404
Adjusted R2	0.384
Note:	*p<0.1; **p<0.05; ***p<0.01

Appendix G Table 2: Secondary analysis

	Dependent variable: CMAS-UK average score
Treatment	0.004 (-0.014)
Constant	2.104*** (-0.082)
Observations	2,181
R2	0.119
Adjusted R2	0.089
Note:	*p<0.1; **p<0.05; ***p<0.01

## Appendix H: Output of missing data analysis

Appendix H Table 1: Missingness of outcome data by treatment arm (randomised sample)

	Comparison (non-pre-teaching)		Pre-teaching	
	<i>n</i> randomised = 1,305		<i>n</i> randomised = 1,304	
	<i>n</i> missing	% missing	<i>n</i> missing	% missing
QRT baseline	65	5	68	5
QRT endline	176	13	201	15
CMAS-UK (endline)	201	15	223	17.1

Appendix H Table 2: Missingness at analysis by treatment arm (randomised sample)

	Comparison (non-pre-teaching)		Pre-teaching	
	<i>n</i>	%	<i>n</i>	%
Complete observations (Primary analysis)	1,107	85	1,079	83
Incomplete observations	198	15	225	17
Withdrawals	123	9	139	11

Appendix H Table 3: Results of missing data regression for covariate missingness

	Dependent variable: Baseline data present
Gender: Male	0.257 (0.208)
FSM: Yes	0.568*** (0.212)
EAL: Yes	0.270 (0.218)
Year group: Year 2	0.029 (0.208)
Treatment	0.052 (0.208)
Constant	-3.636*** (0.247)
Observations	2,411
Note:	*p<0.1; **p<0.05; ***p<0.01

Appendix H Table 4: Results of missing data regression for outcome missingness

	Dependent variable: Outcome data present
QRT_Base_Total_Score	-0.016 (0.016)
Gender: Male	0.193 (0.139)
FSM: Yes	0.218 (0.149)
EAL: Yes	-0.094 (0.154)
Year group: Year 2	-0.328** (0.140)
Treatment	0.177 (0.140)
Constant	-2.142*** (0.214)
Observations	2,314
Note:	*p<0.1; **p<0.05; ***p<0.01

## Appendix I: Regression outputs and effect sizes for subgroup (FSM) analyses

Appendix I Table 1: Regression outputs and effect sizes for subgroup (FSM) analyses

	Dependent variable: Standardised QRT End Total Score	
	Model 4a (FSM only)	Model 4b (Full sample)
Treatment	-0.072 (-0.067)	0.028 (-0.041)
FSM status: Yes		-0.01 (-0.057)
Standardised QRT baseline score	0.470*** (-0.043)	0.548*** (-0.021)
Treatment* FSM		-0.095 (-0.078)
Constant	0.702*** (-0.085)	-0.09 (-0.1)
School fixed effects	Yes	Yes
Observations	610	2,108
R2	0.426	0.401
Adjusted R2	0.351	0.379
Note:	*p<0.1; **p<0.05; ***p<0.01	

## Appendix J: Regression outputs for additional analyses and robustness checks

Appendix J Table 1: Primary and secondary analysis with additional covariates

	Dependent variable:	
	Standardised QRT endline score	CMAS-UK score
	Primary	Secondary
Treatment	0.01 (-0.034)	0.001 (-0.014)
Standardised QRT baseline score	0.544*** (-0.02)	
EAL: Yes	-0.043 (-0.047)	-0.013 (-0.019)
Year group: Year 2	0.006 (-0.035)	-0.009 (-0.015)
Gender: Male	-0.015 (-0.035)	-0.035** (-0.015)
Constant	-0.074 (-0.1)	2.129*** (-0.083)
School fixed effects	Yes	Yes
Observations	2,145	2,139
R2	0.405	0.12
Adjusted R2	0.384	0.089
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	

Appendix J Table 2: Secondary analysis with baseline QRT score

	Dependent variable: CMAS-UK average score
Treatment	0.005 (-0.014)
Standardised QRT baseline score	-0.019** (-0.008)
Constant	2.105*** (-0.085)
Observations	2,135
R2	0.120
Adjusted R2	0.089
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Appendix J Table 3: Primary analysis with 'on time'/'late' indicator

	Dependent variable: QRT end total score
Treatment	-0.051 (0.051)
On time: 1	-0.241 (0.149)
Standardised QRT baseline score	0.544 (0.020)
Treatment* 'on time'	0.103 (0.068)
Constant	-0.056 (-0.099)
School fixed effects	Yes
Observations	2,186
R2	0.405
Adjusted R2	0.384
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Appendix J Table 4: FSM (interaction) analysis with additional covariates

	Dependent variable: QRT end total score
Treatment	0.03 (-0.041)
FSM: Yes	-0.018 (-0.058)
Standardised QRT baseline score	0.547*** (-0.021)
EAL: Yes	-0.04 (-0.048)
Year group: Year 2	0.021 (-0.036)
Gender: Male	-0.012 (-0.036)
Treatment* FSM	-0.096 (-0.078)
Constant	-0.088 (-0.103)
School fixed effects	Yes
Observations	2,083
R2	0.402
Adjusted R2	0.38
Note:	*p<0.1; **p<0.05; ***p<0.01

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0.

To view this licence, visit <https://nationalarchives.gov.uk/doc/open-government-licence/version/3> or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk)

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.


This document is available for download at <https://educationendowmentfoundation.org.uk>



Education  
Endowment  
Foundation

The Education Endowment Foundation  
5th Floor, Millbank Tower  
21–24 Millbank  
London  
SW1P 4QP

<https://educationendowmentfoundation.org.uk>

 @EducEndowFoundn

 [Facebook.com/EducEndowFoundn](https://www.facebook.com/EducEndowFoundn)