Same Day Intervention

Evaluation Report

November 2020
Updated: April 2021

Malen Davies, Shivonne Gates, Anysia Nguyen, Robert Wishart and Daniel Phillips

NatCen
Social Research that works for society

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus – Private Equity Foundation) and received a founding £125m grant from the Department for Education. Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

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

Jonathan Kay
Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

0207 802 1653

jonathan.kay@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator 3

Executive Summary 4

Introduction 6

Methods 13

Impact evaluation results 28

Implementation and Process Evaluation results 40

Conclusion 60

References 64

Appendix A EEF cost rating 65

Appendix B Security classification of trial findings 66

Appendix C Impact evaluation figures 67

Appendix D Effect size estimation 70

Appendix E Intervention description 72

Further appendices 73

About the evaluator

The project was independently evaluated by a team from NatCen Social Research: Daniel Phillips, Malen Davies, Shivonne Gates, Anysia Nguyen and Robert Wishart.

Contact details:

Daniel Phillips
NatCen Social Research
35 Northampton Square
London EC1V 0AX
0207 549 7184
daniel.phillips@natcen.ac.uk

Executive Summary

The project

Same Day Intervention (SDI) is designed to increase pupils' mathematics attainment across all age groups and provide targeted support so that all pupils attain a given level of mathematics understanding. Business-as-usual mathematics lessons are replaced with approximately 75-minute SDI lessons, which are separated into three parts. In the first 30 minutes, teachers model a mathematics concept such as subtraction and pupils complete five to six diagnostic questions to assess their understanding of the concept. In the following 15 minutes, a 'progress pit-stop' takes place, where teachers mark pupil responses while pupils attend an assembly or are taught by a Teaching Assistant (TA). In the final 30 minutes, the teacher uses the assessment to separate the class into two, providing additional instruction to an 'SDI' group of pupils who did not understand the concept, while a TA supports the rest of the class as they engage with more challenging content. The programme aims to improve the mathematics attainment of all pupils but is specifically designed to support lower attaining pupils in understanding key mathematics concepts. It also aims to reduce teacher workload by providing teachers with training and resources and reducing their marking workload outside of the classroom. In this trial, Outwood Institute of Education trained teachers, Senior Leadership Team (SLT) staff and TAs through a combination of training sessions and follow-on support, while lesson resources were also provided by the Yorkshire and the Humber MathsHub. Teachers were then expected to deliver daily SDI mathematics lessons, with the support of a TA, to Year 5 pupils (aged nine to ten) for seven months.

The evaluation was a two-armed randomised controlled efficacy trial. In total, 73 schools (3,298 pupils) from Yorkshire and the Humber and surrounding areas participated, with 37 schools in the intervention group and 36 schools in the control group. The process evaluation included ten case studies of intervention schools, which involved in-depth teacher interviews and pupil focus groups. Telephone interviews were also conducted with five control schools, alongside an interview with the developer, and surveys of schools and teachers. School recruitment for the trial began in January 2018 and post-testing was completed in July 2019.

Key conclusions
1. Pupils in Same Day Intervention schools made the equivalent of zero additional months' progress in mathematics, on average, compared to pupils in control schools. This result has a moderate to high EEF security rating.
2. Half of teachers in Same Day Intervention schools provided data on the time they spent marking pupil work, and there was evidence that they spent less time marking than teachers in control schools (approximately one hour less per week). Teachers in Same Day Intervention schools were also more likely than teachers in control schools to report that all pupils, and lower achieving pupils, were more confident compared to the previous year's pupils.
3. Exploratory analysis suggests that pupils eligible for Free School Meals (FSM) in Same Day Intervention schools made the equivalent of zero months' progress in mathematics, on average, compared to FSM pupils in control schools.
4. Fewer than half of intervention schools implemented Same Day Intervention as intended. Schools faced several challenges which made implementation difficult, such as being able to provide a teaching assistant to support delivery and adapting the school day to extend mathematics lessons.
5. Compliance analysis suggests that pupils in schools that implemented Same Day Intervention as the programme developer intended may have made, on average, two months additional progress in mathematics compared to pupils in control schools. There may be other differences in schools that implemented the programme with greater fidelity, however, making this finding less secure than the primary analysis.

EEF security rating

These findings have a moderate to high security rating. This was an efficacy trial, which tested whether the intervention worked under developer-led conditions in a number of schools. The trial was a two-armed randomised controlled trial. The trial was not as well-powered as originally intended because the number of schools recruited was lower than expected. Of the pupils who started the trial, 14.2% were not included in the final analysis because they did not sit the post-test or could not be matched to the National Pupil Database (NPD). There were some important differences in prior attainment between the pupils in Same Day Intervention schools and those in comparison schools. The presence of a moderate risk of control schools doing similar interventions, and variable implementation fidelity, make it harder to accurately estimate the size of the impact on the pupils in the trial.

Additional findings

Pupils in Same Day Intervention (SDI schools made, on average, zero additional months' progress compared to those in the control equivalent. This is our best estimate of impact, which has a moderate to high EEF security rating. As with any study, there is always some uncertainty around the result; the possible impact of this programme also includes negative effects of up to two months' less progress and positive effects of up to two additional months' progress. Exploratory analysis suggests that children eligible for Free School Meals in SDI schools also made the equivalent of zero additional months' progress in mathematics, on average, compared to FSM children in control schools (with the possible range of effects ranging from three months' less progress to two additional months' progress).

In terms of secondary outcomes, half of teachers in SDI schools provided data on the time they spent marking pupil work. There was evidence that they spent less time marking than teachers in control schools (approximately one hour less per week). However, given that data is only available for half of teachers (due to teacher turnover and survey non-response), this should be treated with caution. Findings from qualitative case studies, and pre- and post-intervention surveys from the Implementation and Process Evaluation (IPE), also suggested that marking time may have decreased. However, they also indicated that the time taken to plan lessons may have increased due to the variety of resources that teachers needed to plan for each component of the lesson. While some teachers suggested that planning time reduced throughout the year, others reported that planning workload remained higher than usual throughout the programme. Another secondary outcome was teachers' perception of pupils' confidence. Teachers in SDI schools were more likely than teachers in control schools to report that all pupils, and lower achieving pupils, were more confident compared to the previous year's pupils. The Implementation and Process Evaluation corroborated this finding, with teachers reporting that the change in structure of mathematics lessons helped to build the resilience, confidence, and self-esteem of pupils. It should be noted that this is the perception of teachers, rather than pupils. Pupil focus groups in case study schools suggested that pupils had mixed feelings towards SDI.

Implementation analysis indicated that fewer than half of intervention schools implemented SDI as intended, and schools faced several challenges which made effective implementation difficult. For instance, schools were required to ensure a TA was present in every Year 5 mathematics lesson, as their support was a core component of the intervention. The post-intervention teacher survey found that 25% of schools always had a TA present, 46% regularly did, 18% only occasionally did, and 9% never did. Schools also found it challenging to adapt the school day so that sessions could last the intended 75 minutes. Other aspects of delivery (such as the approach to modelling concepts, the diagnostic test and accompanying pit-stop to mark assessments, and the division of the class into targeted groups) were also often not implemented with fidelity, while attendance at training was also mixed. Only 55% of teachers attended all three compulsory training sessions. Surveys and case studies suggest that the lack of TA presence, regular teacher absence due to other training commitments, the need to meet other whole-school learning objectives set by SLT, and a lack of buy-in from some teachers into SDI made it difficult for schools to overcome these implementation challenges.

Poor implementation may have contributed to the overall lack of impact. Compliance analysis suggests that when SDI was implemented with greater fidelity pupils made, on average, two months' additional progress in mathematics (with the possible range of effects ranging from no additional progress to four months' additional progress). Other factors may also have contributed to the lack of impact, including the programme not being taught for long enough, or control schools using similar practices. This study represents the first robust evaluation of the SDI programme. It is unable to replicate the promising finding from the pilot study of SDI (McGarry, 2017); a study which had methodological limitations. Broader prior evidence indicates that interventions that adopt a similar model can influence mathematics attainment, but do not always do so, a finding confirmed here.

Cost

The average cost of SDI for one school was around £13.94 per pupil per year when averaged over three years. The intervention also requires approximately 130 days of school staff time to deliver SDI, in addition to 74 school staff days spent preparing outside of school, alongside five days of training, and six days of lesson cover.

Table 1: Summary of impact on primary outcome

Outcome/ Group	Effect size [95% confidence interval]	Estimated months' progress	EEF security rating	Number of pupils	p-value	EEF cost rating
Mathematics attainment: full sample	0.00 [-0.15; 0.16]	0	🔒🔒🔒🔒🔒	2,830	0.698	£ £ £ £ £
Mathematics attainment: have ever been eligible for FSM (subsample model)	-0.02 [-0.22; 0.17]	0	N/A	836	0.578	N/A

Introduction

Rationale for intervention

Same Day Intervention (SDI) has its origins in the England-China Mathematics Teacher Exchange (MTE) of 2014–2015, a programme implemented by the National Centre for Excellence in the Teaching of Mathematics (NCETM) and the national MathsHub Programme. The England-China MTE involved teachers and school leaders from the UK visiting Shanghai to observe mathematics classes, and then hosting Chinese teachers in the UK.

SDI is intended to adapt elements of the Shanghai model of mathematics teaching for the UK context. It takes its name from a key element of the intervention: after teaching a new mathematics concept, teachers assess pupils' understanding using diagnostic questions, separate the class into groups and then lead interventions designed to ensure that all pupils grasp the key elements of that concept by the end of a class. Many of the ideas and concepts from the Shanghai model of mathematics teaching (and those that underpin SDI) have been implemented in UK classrooms prior to this evaluation, in various schools within the MathsHub Programme, the Outwood Grange Academies Trust (OGAT), and beyond. The MathsHub Programme advocates nationally for adoption of key elements of the Shanghai model of teaching, including SDI (NCETM, 2015).

However, there is little programme-level evidence exploring the effectiveness of this type of intervention in mathematics. An evaluation conducted by McGarry (2017) explored implementation of a pilot of this type of intervention in several schools in the north of England, finding that pupils' attitude towards mathematics and their attainment in mathematics increased when pre-intervention and post-intervention scores were compared. However, this evaluation was small scale and did not seek to compare an intervention group with a group of schools that were not implementing the programme.

There is a wider body of evidence exploring the effectiveness of some key features of the SDI approach. A recent systematic review (Hodgen *et al.,* 2018) provides evidence on the effectiveness of 'mastery learning' for mathematics in which teachers adopt a variety of approaches to teaching, including frequent feedback and additional time for struggling students—a model of teaching that shares several aspects in common with SDI's learning strategies, embedded feedback, and targeted teaching. Hodgen *et al.* (2018) find that mastery learning can be effective in increasing mathematics attainment, but this conclusion draws largely on an evidence base from the US, with the UK evidence base being far smaller and indicating that, at best, effects on mathematics attainment are small. Hodgen *et al.* (2018) also find evidence that providing feedback to pupils (as SDI does through an assessment and subsequent group tasks in each class) can have a large effect on learning, but that the range of effects is quite broad, and some studies show that it can have negative effects. SDI's assessment and group tasks also fit with some recommendations on how to make feedback effective, such as ensuring that it is task-related and designed to encourage effort (Hattie & Timperley, 2007), or it is used specifically for conceptual tasks (Soderstrom & Bjork, 2013). However, there is also evidence that feedback may be most effective where it is used sparingly (Soderstrom & Bjork, 2013). Evidence on the effectiveness of within-class attainment grouping (where, as with SDI, pupils are organised within their class for some specific activities) indicates that it is likely to benefit all learners but that any impact may be lower for lower attaining pupils than for others (EEF, 2018a). EEF (2018a) also finds that use of TAs in-class to support small groups has, on average, moderate positive benefits.

This efficacy trial was designed to evaluate whether SDI has an impact on pupil attainment and intermediate outcomes such as teacher marking workload and teachers' perceptions of their pupils' confidence in mathematics. The trial was designed as a two-armed, four-level cluster randomised controlled efficacy trial. The four levels included pupils within classes, within schools, and within three regional hubs. Randomisation was conducted at school level to minimise spill-over effects which could have biased effect estimates. Randomisation was stratified by regional hub to allow for any differences in implementation and school characteristics between the three regional hubs. Schools allocated to the control condition carried on teaching mathematics as usual, while schools in the intervention condition implemented SDI mathematics classes. The Implementation and Process Evaluation (IPE) was designed to understand and explore how SDI was delivered, whether there were any barriers or facilitators to delivery and what the context underpinning the delivery of the intervention was.

Intervention

Overview:

SDI classes are taught daily in place of daily business-as-usual mathematics classes. Teachers start the class by modelling a mathematics concept such as subtraction. They employ 'teach-practice' or ping-pong style teaching (where teachers model the concept being taught before asking pupils to complete a related task, with support from the teacher), supported by Assessment for Learning strategies, scaffolding, and differentiated questioning. Pupils then independently complete five or six diagnostic questions designed to test their understanding. There is then a 15-minute 'progress pit-stop' during which teachers mark pupils' work and group them based on their task performance. During this time, pupils may attend an assembly or be taught by a Teaching Assistant (TA). Pupils who need further help to fully understand the concept form the 'SDI group'. The teacher then provides the SDI group with further instruction with the aim of addressing common misconceptions and embedding learning. Pupils who demonstrate that they have already grasped a concept are given more challenging tasks, designed to deepen and master the topic, supported by the TA. The SDI lesson is designed to last 75 minutes with two periods of roughly 30 minutes in length sandwiching the 15-minute progress pit-stop. To be implemented as intended, the intervention requires a TA to be available for all SDI classes in order to supervise pupils during the progress pit-stop and to support those pupils who are not in the SDI group. SDI is designed for pupils in Key Stage 1 or 2, but this trial explored the impact of SDI for pupils in Year 5.

For a diagrammatic overview of the SDI approach, see Appendix E.

Why:

SDI is designed to increase pupils' mathematics attainment and provide targeted support so that all pupils attain a given level of mathematics understanding and thus prevent an achievement gap from emerging.

What: Physical or informational materials used in the intervention

Teachers could source or write their own diagnostic questions and mark pupils' work according to a 'marking for diagnosis code' which was developed by the Yorkshire and the Humber MathsHub. The marking for diagnosis code, a bank of questions, and exemplar lesson resources were provided during training and made available online and on google classroom or Padlet,[1] with teachers also encouraged to use these to produce their own resources over the course of SDI implementation.

What: Procedures, activities, and/or processes used in the intervention

Staff from all intervention schools were required to attend three training sessions delivered by the Outwood Institute of Education (OIE). The first training session was an Open Classroom (OC) where up to 20 teachers and Senior Leadership Team (SLT) staff were able to observe an SDI lesson, receive initial training on how to plan a lesson and discuss the OC. The OCs were delivered by 'host' schools, which were already delivering SDI across all year groups within their school. The intention for the OC sessions was for teachers and SLT staff to see SDI in action with Year 5 pupils. A total of three OC sessions took place to ensure that pupils were not being observed by too many people at once. During the OC, teachers and SLT staff suggested to the developer that it would be useful for their TA to observe an OC SDI lesson. In response to this, the developer also invited all TAs to an OC training session. TAs who could not attend the OC training had the training cascaded down to them by teachers.

Teachers also attended a further two days of training focusing on the different skills needed to deliver SDI. Training day two focused on assessing the diagnostic task given to pupils, using the marking for diagnosis code and leading the SDI component of the lesson. Training day three focused on designing activities for the second half of the lesson (known in the training as the Bronze, Silver, and Gold tasks) which pupils complete with support from the TA. Teachers and SLT staff were also invited to attend surgery sessions throughout the year. Their purpose was to offer a space for teachers and SLT staff to discuss delivery and 'troubleshoot' any issues they were experiencing. Ongoing support was also

[1]Padelt was a platform used to deposit lesson resources for SDI teachers. See Further appendix 9 for examples of these resources.

provided in the form of online resources posted on an online portal called Padlet and telephone support from the developer.

Table 2 provides a breakdown of how each component of the intervention was intended to be delivered during the trial:

Table 2: SDI components

Component		Time needed
Whole-class input	Includes 'teach-practice' techniques, Assessment for Learning strategies, scaffolding and differentiated questioning.	Approximately 30 minutes for whole class input and diagnostic
Diagnostic assessment	Five or six carefully crafted questions are completed independently by pupils.	
Pit-stop	The teacher marks pupils' diagnostic assessment and groups the pupils based on their assessment. During this time pupils can be at an assembly or taught by the TA.	15 minutes
SDI grouping	The diagnostic task is used to group pupils. Those who demonstrate that they need further help to grasp a concept form the 'SDI group' and are taught by the teacher. Pupils who demonstrate that they have grasped a concept are given more challenging tasks designed to deepen and master a topic, supported by the TA.	Approximately 30 minutes
Teaching Assistant (TA)	Availability of a TA for all SDI classes.	

Who: Intervention providers/implementers

Teachers implement the intervention with the help of TAs. SDI is designed so that the teacher leads the class and provides additional support to those who need it following the diagnostic task (the 'SDI group'), while TAs support pupils who did well on the diagnostic task and have already shown that they have grasped the concept being taught. In addition, SLT staff help create the structural change to the school day.

How: Mode of delivery

In this trial, SDI classes replaced traditional mathematics classes and were intended to be delivered daily to Year 5 pupils in primary schools in England. While pupils still received a 60-minute mathematics lesson, there was a change to the timetable so that the teacher could be free for an additional 15 minutes for a progress pit-stop to mark a diagnostic test. Thus, SDI classes were 15 minutes longer than usual, taking place over 75 minutes. This also meant a change to the way a TA worked with the class (supporting the pupils who had grasped the concept being taught and possibly leading a session for the whole class during the progress pit-stop).

Where: Location of the intervention

SDI was delivered to non-selective primary schools in Yorkshire and the Humber and surrounding areas. Training was delivered separately for each of three regional hubs, Lincolnshire, Derbyshire, and Oldham.

When and how much: Duration and dosage of the intervention

SDI is usually designed to be a continuous practice over the school year and can be implemented across school Years 1 to 6. In this trial, it was designed to take place in regular mathematics classes over approximately nine months but, as school recruitment was completed later than anticipated, implementation actually took place over approximately seven months. SDI classes last 75 minutes and are designed to be split into two 30-minute periods with a 15-minute progress pit-stop in the middle of them.

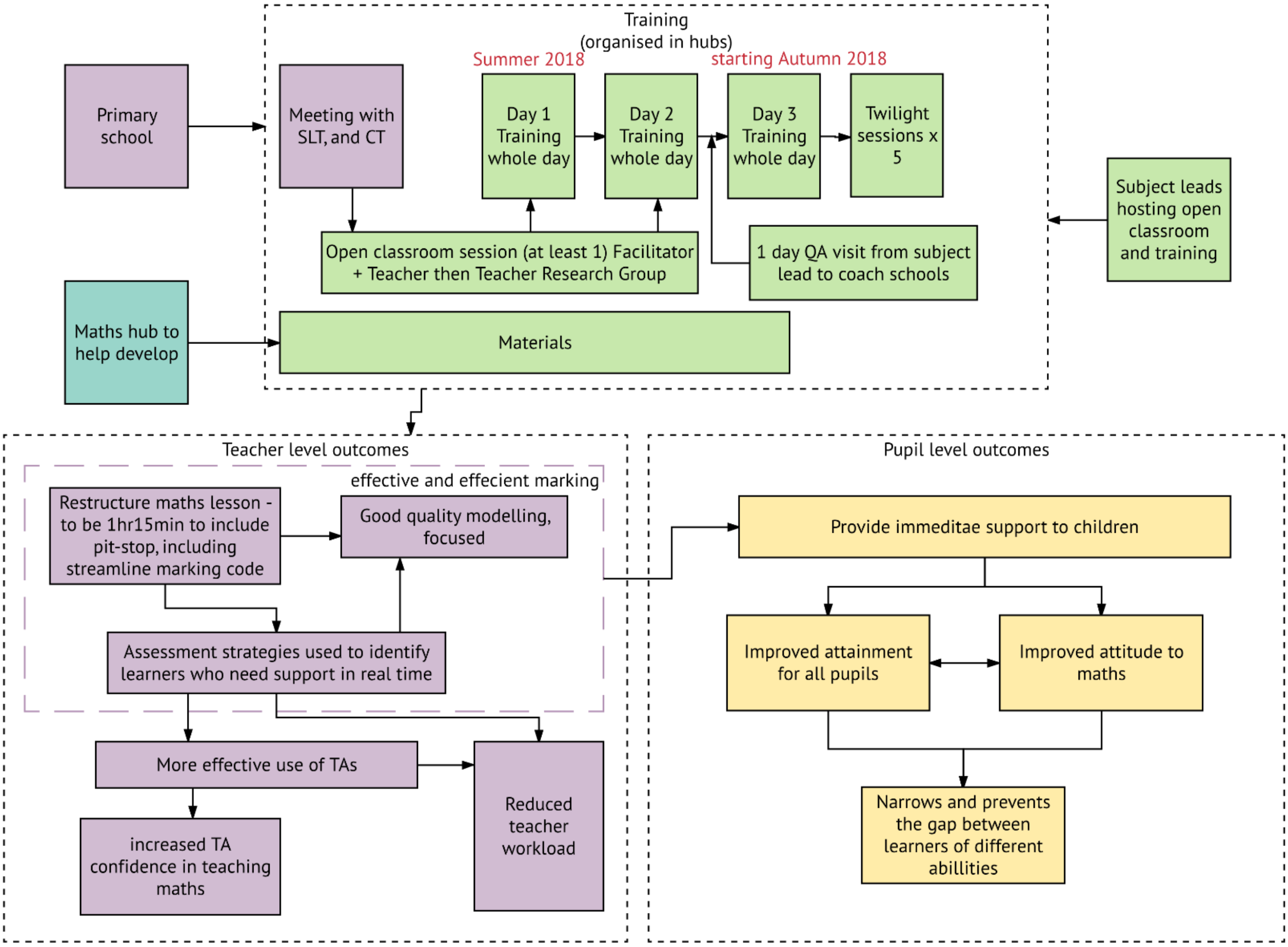
Same Day Intervention Theory of Change

Figure 1 presents the SDI Theory of Change (ToC), as specified in the study protocol. The programme involves training teachers and TAs in how to deliver SDI and requires that a representative from a school's SLT attends one training session in order to promote buy-in to the intervention and help provide support for the structural changes to mathematics lessons that SDI requires.

The ToC assumes that staff attend training and will buy-in to the programme, implement its ideas and structure, and make use of supporting resources provided on the online platform or Padlet. It also assumes that a TA be made available for all SDI classes and that structural changes to the school day (an extended mathematics class of 75 minutes including a 15-minute pit-stop) will be made. SDI classes should replace all regular mathematics classes.

The programme is designed to reduce teacher workload by providing them with training, resources, and a new and more effective approach to teaching mathematics. It is intended to reduce teachers' workload outside of class by allowing them to mark pupils' work as part of the progress pit-stop. It is designed to promote pupil confidence in mathematics by ensuring that all children receive the support they need to understand each mathematics concept, before lessons progress to a new concept. By tailoring support to pupils' needs, SDI aims to increase mathematics attainment for all pupils and to close any achievement gap or prevent one from emerging. Finally, the ToC hypothesises that SDI may increase TA confidence in teaching by introducing them to some new teaching techniques. An updated ToC and accompanying summary is presented in the IPE section of this report.

Figure 1: Same Day Intervention Theory of Change



Overview of trial recruitment, implementation and testing

The evaluation of SDI was originally designed to incorporate 120 schools. Recruiters approached 144 schools, expecting around 20% would not participate. Of the 144 schools approached, 95 schools agreed to participate. Among those that agreed to participate, 22 schools did not complete all tasks necessary to fully sign-up to the intervention and evaluation (sign a Memorandum of Understanding [MoU], enumerate pupils, and complete baseline surveys). Therefore, only 73 schools were successfully recruited and randomised for the efficacy trial. It may be that the completion of sign-up tasks was too demanding for school staff or that the requirements of the intervention as set out in the MoU (to extend the length of mathematics classes and to make a TA available for each SDI lesson) put off some schools. We did not collect data from schools that did not take part in the evaluation and therefore could not assess if they systematically differed to those that took part. The sample size calculation section provides further details on power calculations as envisaged at protocol, randomisation, and analysis.

The intention was for SDI training to happen in the Summer term, with teachers beginning the intervention from September. Due to the recruitment challenges outlined above, training was moved to September to allow for a longer recruitment period. This meant that teachers had seven months to deliver the intervention rather than the originally intended nine months.

The compulsory SDI training received mixed attendance from teachers and representatives of schools' SLT. Of the treatment teachers 91% attended at least one of the three sessions (although only 55% attended all three sessions), while 60% of schools sent an SLT representative to attend training. However, there were some issues with implementation fidelity in schools. For example, not all schools were able to provide TAs for all classes. Furthermore, of the 36 intervention school teachers who responded to the survey, 20% (n=11) were not consistently using the SDI components in mathematics classes, while 20% of teachers (n=11) reported only including the pit-stop occasionally and 7% (n=4) did not include it in any lessons. SDI classes were not always extended so that they lasted the required 75 minutes. In addition, schools did not always implement the key elements which are: the whole-class input or approach to modelling mathematics concepts, the diagnostic test of pupils' understanding and accompanying pit-stop to mark pupils' tests, and the division of the class into targeted groups according to the result of their diagnostic test. Finally, not all schools implemented SDI for the anticipated period of approximately seven months. We explore these issues in full in the IPE element of this report.

While 73 schools (3,298 pupils) were randomised to intervention and control conditions, one school withdrew from the trial before endline testing could take place. The final sample was therefore 72 schools (3,238 pupils). In total, 2,892 pupils sat the endline test, GL Assessment's *Progress Test in Maths 10* (*PTM10*). The remaining 346 pupils did not sit the test due to being on holiday, ill, having left school, or were pupils with Special Educational Needs and Disabilities (SEND).[2] Furthermore, 62 pupils could not be linked to the National Pupil Database (NPD) that provided our baseline measure of attainment, leaving a total of 2,830 pupils in 72 schools for the primary outcome analysis. This means that total rate of pupil attrition from randomisation to analysis is 14.2% (15.5% in intervention schools and 13.1% in control schools).

Evaluation objectives

The impact evaluation addresses the following research questions:

Primary research question:

- What is the impact of SDI on Year 5 pupils' mathematics attainment in non-selective state schools in England?

Secondary research questions:

- To what extent does participation in SDI affect teacher marking workload?
- To what extent does participation in SDI affect Year 5 teachers' perceptions of their pupils' confidence in their mathematics abilities?
- Does the impact of SDI on Year 5 pupils' mathematics attainment differ by FSM eligibility?[3]
- What is the impact of SDI on the size of the gap between higher and lower achieving Year 5 pupils?

The implementation and process evaluation addresses the following research questions:

1. How was the intervention delivered across the range of delivery sites?
2. To what extent did contextual variation affect fidelity? What adaptations were put in place?

---

[2]NatCen asked that every pupil who would usually sit a national test like SATS, sit our follow-up test. When children's needs or disabilities were too severe, teachers alerted us that these pupils would not sit the test.
[3]Following EEF's guidance (2018), the subgroup analysis by FSM eligibility was conducted as part of the secondary analysis, rather than the primary analysis as previously outlined in the protocol (Davies *et al.*, 2019). In addition, this was assessed with both an interaction model (as specified in the Statistical Analysis Plan [SAP]) and as a subsample (as required by EEF's guidance).

3. What were the barriers to delivery and how were these addressed by schools and trainers?
4. What facilitated successful delivery?
5. To what extent did participation in SDI affect teachers' confidence and workload?
6. To what extent did participation in SDI affect teacher perceptions of pupils' attitudes towards mathematics, their confidence, and self-efficacy?[4]
7. What is the 'business as usual' approach to teaching of Year 5 mathematics in control schools?[5]

You can access the study **protocol** and **statistical analysis plan** (SAP) by clicking on the relevant link.

Ethics and trial registration

Ethical approval to carry out the study was granted by NatCen's Research Ethics Committee (REC). The trial was registered on the **ISRCTN Registry** with trial registration number 43822826.

Each school that expressed an interest in signing up to the intervention was sent a MoU (please see a copy in the Further appendix 3). This MoU covered a commitment to deliver the intervention in the event of allocation to the intervention group and to support evaluation activities.

Data protection

NatCen was the data controller and data processor for the project, meaning that NatCen was responsible for deciding the legal basis for processing data. In line with the General Data Protection Regulation (GDPR) Article 6, 'legitimate interest' was chosen as the legal basis for processing personal data. Legitimate interest was selected on the basis that establishing whether SDI is a successful intervention is of benefit to primary schools in England.

NatCen had access to pupil, school and teacher information, recordings of interviews, transcripts, charted data, sample files (including contact details), survey responses, and test results. All responses were anonymised before being analysed.

The central purpose for processing data was to evaluate the effectiveness of the SDI. All results collected were used for evaluation purposes only. Data gathered from interviews, surveys, and focus groups was used to inform the process study element of the project, while pupils' mathematics attainment from GL Assessment's *PTM10* were used for the impact study element.

Third parties who would also have access to certain information were listed as the intervention developers at the OIE; McGowan transcription; Fomara Print+; and GL Assessment.

At the end of the research, all pupil data will be anonymised before being archived. All personal information, and any other data held on the project, will be securely deleted once the project is complete in April 2021.

NatCen is the data controller and processor from evaluation inception, up until relevant data is archived and securely deleted. EEF will become the data controller once the datasets are archived.

A copy of the privacy statement published on the NatCen website can be found in the Further appendix 1.

Project team

The developers at Yorkshire and the Humber MathsHub, which is led by the Executive Director of Mathematics for OGAT and the OIE were responsible for overseeing the delivery of SDI in schools.

The evaluation of SDI was led by NatCen's Children and Families team, who were responsible for the day-to-day project management of the study. The team worked closely with impact evaluation experts from NatCen's Evaluation team, who provided expertise on the design of the trial and testing. Contact with schools, pupil testing, and data entry

---

[4]Question re-phrased from that originally contained in the study protocol (Davies *et al.*, 2019) which was 'To what extent does participation in SDI affect pupil attitudes towards maths, their confidence and self-efficacy?'
[5]Question omitted from study protocol (Davies *et al.*, 2019).

processes were coordinated through NatCen's Operations department. The analysis plan was written by Daniel Phillips, Robert Wishart, Malen Davies and Anysia Nguyen. Members of the project delivery team, Rakhee Patel and Tanya Basi moved onto other organisations during different stages of the project.

Table 3: Project team

Project team		
Daniel Phillips	Research Director	Overall study lead
Malen Davies	Research Director	Project lead
Robert Wishart	Research Director	Impact evaluation
Shivonne Gates	Senior Researcher	IPE and testing support
Anysia Nguyen	Researcher	Testing and impact evaluation support

Methods

The report presents the results of the evaluation of a two-armed, four-level cluster randomised controlled efficacy trial. The highest level of clustering was regional hub (fixed effect). Nested within regional hubs were three levels: schools, classes, and pupils (random effects). Randomisation was conducted at school level to minimise spill-over effects which could have biased effect estimates. Schools allocated to the control condition carried on teaching as usual, whilst schools in the intervention condition implemented SDI classes instead of normal mathematics classes. The main treatment effect was calculated on an intention-to-treat (ITT) basis, assessing the impact of the SDI on Year 5 pupils' mathematics attainment. Table 4 summarises the main characteristics of this efficacy trial.

Table 4: Trial design

Trial design, including number of arms		Two-armed, four-level cluster randomised controlled efficacy trial with random allocation at school level
Unit of randomisation		School
Stratification variable (s) (if applicable)		Regional hub (Lincolnshire, Derbyshire, and Oldham)
Primary outcome	Variable	Year 5 mathematics attainment
	Measure (instrument, scale, source)	GL Assessment's *Progress Test in Maths 10* (*PTM10*) (scale: 0–65)
Secondary outcome(s)	Variable(s)	Teacher marking workload in class time and outside of class in an average week, and teachers' perceptions of pupil confidence (overall and lower achievers)
	Measure(s) (instrument, scale, source)	Teacher survey
Baseline for primary outcome	Variable	EYFSP and KS1 mathematics attainment
	Measure (instrument, scale, source)	EYFSP mathematics 'Numbers' and 'Shape, space and measures', and KS1 mathematics outcome
Baseline for secondary outcome(s)	Variable(s)	Teachers' workload, and teachers' perceptions of pupil confidence
	Measure (instrument, scale, source)	Teacher survey

The primary outcome of interest was Year 5 pupils' mathematics attainment as measured by *PTM10*. The SAP (Phillips *et al.*, 2019) outlined the derivation of a baseline attainment measure, combining the Early Years Foundation Stage Profile (EYFSP) mathematics Numbers and Shape, space and measures scores with Key Stage 1 (KS1) mathematics scores. However, due to the lack of variance in this derived measure, the primary analyses used solely KS1 mathematics scores as a measure of prior attainment. An additional sensitivity analysis was conducted, including KS1 and EYFSP scores separately in the regression model (see Table 19).

Secondary outcomes included teacher marking workload (time spent marking in and out of class on an average week), teachers' perception of pupil confidence overall, and teachers' perception of confidence for low achievers. This was measured using a bespoke survey created by NatCen (see Further appendix 2).

Participant selection

Schools were judged eligible to take part if they met the following criteria:

1. Location – schools had to be based in Yorkshire and the Humber and surrounding areas;
2. Number of classes per school – schools needed to have at least one class of Year 5 pupils.[6] Mixed year-group classes were not eligible;
3. SDI involvement – eligible schools could not have been involved in any SDI projects previously, although they could have employed other Shanghai methodologies or been involved in a MathsHub project;
4. School type – only mainstream non-selective and non-special schools were eligible.

All pupils in Year 5 classes in schools that met the four criteria above were eligible. were recruited by the Yorkshire and the Humber MathsHub. MathsHubs bring together a network of schools and colleges driving improvements to mathematics education, led by the National Centre for Excellence in the Teaching of Mathematics (NCETM and funded by the Department for Education. The Yorkshire and the Humber MathsHub is led by the Executive Director of Mathematics for OGAT and the OIE. It delivers SDI training and provides support and access to teaching resources. They were drawn from the population of schools in Yorkshire and the Humber and surrounding areas. The Yorkshire and the Humber MathsHub and OIE delivered several information sessions in the local area to showcase the intervention and trial. Marketing materials were also developed, including a dedicated webpage and leaflets. The webpage allowed schools to submit their interest in the trial. These schools were followed up and sent a Memorandum of Understanding (MoU to be signed by SLT staff, the Yorkshire and the Humber MathsHub, and NatCen. The MoU provided an overview of the intervention and detailed requirements for the three parties,[7] and also collected consent to participate in the study. The full MoU can be found in the Further appendix 3.

Once participating schools had been identified, schools sent a letter to parents of Year 5 pupils to inform them about the schools' participation in the trial and programme. The letter explained that schools' access to the programme would be determined by lottery. It also provided the opportunity for parents to withdraw from having their child's data included in the analysis or any other aspects of the trial. The letter sent to parents can be found in Further appendix 4.

Originally, the trial of the SDI was designed to incorporate 120 schools to achieve a minimum detectable effect size (MDES) of 0.205. However due to reasons outlined in the 'Overview of trial recruitment, implementation and testing' section, only 73 schools were successfully recruited and randomised to the SDI trial. Consequently, the trial was not sufficiently powered to detect small treatment effects. One school withdrew from the trial mid-year, before testing at follow-up. Therefore, the final analytical sample contained 72 schools and 3,238 pupils: 1,522 in the intervention group and 1,716 in the control group. The 'Sample size' section provides further details on power calculations as envisaged at protocol, randomisation, and analysis. See Figure 2 in the 'Impact evaluation results' section for the full CONSORT diagram.

Outcome measures

Baseline measures

Baseline measures of attainment were sourced from the National Pupil Database (NPD. As outlined in the SAP (Phillips *et al.*, 2019), there are clear limitations to using KS1 and EYFSP scores. As they are both categorical measures, the amount of variance is limited. An alternative would have been to conduct testing at baseline. However, in addition to time and budget constraints, this would have placed additional burden on participating schools and pupils. Not all pupils were successfully linked to the NPD.[8] Of the 3,238 pupils randomised, 3,097 pupils were linked to EYFSP data and 3,169 to KS1 data.

---

[6]It was not a requirement for all eligible classes in participating schools to take part in SDI. Schools were asked to specify how many classes would be participating.
[7]The three parties were the school, the MathsHub, and NatCen Social Research.
[8]Failure to link pupil records could have occurred for several reasons, such as incorrect identifying information being provided when pupils were enumerated.

To mitigate the low variance of KS1[9] and EYFSP[10] mathematics scores, a composite index of baseline attainment was derived as a weighted sum of the individual components. Greater weight was placed on KS1 mathematics (66%) than on EYFSP G11 (Numbers, 17%) and EYFSP G12 (Shape, space and measures, 17%). This weighting reflected the fact that KS1 scores are likely to be better predictors as they were collected more recently. However, the distribution of the derived variable was insufficient to be included in the analysis as a continuous variable. The primary analysis, therefore, used KS1 scores as a categorical measure of baseline attainment. To test the sensitivity of this approach, both EYFSP measures were added in our fully adjusted model.

Primary outcome

SDI targeted Year 5 pupils who do not participate in any national tests in mathematics at the end of the school year. Instead, GL Assessment's *Progress Test in Maths* (*PTM*) was used. *PTM* is a standardised assessment of pupils' mathematical skills and knowledge and was chosen on the basis that it represents a validated assessment of mathematics attainment. The *PTM10* test was selected as it is designed for use with Year 5 pupils. The primary outcome for the evaluation was therefore mathematics attainment as measured by the raw score of *PTM10*. *PTM10* is an untimed test that consists of 20 exercises and a 10-minute mental mathematics test. Pupils took the test under exam conditions as a class. *PTM10* was invigilated by NatCen staff who were blinded to whether schools were in the treatment or control group. *PTM10* test papers were collected by NatCen invigilators and marked by GL Assessment. A total of 2,892 out of 3,238 pupils took the test at endline.

In total, NPD records were matched with *PTM10* results for 2,830 pupils (i.e. attrition of 62 cases or a total of 14.2% from the expected analytical sample).

Secondary outcomes

The evaluation had four secondary outcomes:

- Teachers' workload in terms of:
  - overall time spent marking (including time in and outside of class); and
  - time spent marking outside of class.

- Teachers' perception of pupils' confidence
  - overall; and
  - lower achievers.

Teachers' perceptions of pupils' confidence were not measured at baseline because teachers may have begun teaching a new cohort of pupils at the start of the academic year. Teachers were instead asked in the post-intervention survey to judge current pupils' confidence compared to the cohort from the previous academic year (2017/2018). Possible answers were: a lot less confident, a little less confident, about the same, a little more confident, and a lot more confident. They were coded one to five, with higher scores indicating higher confidence (see Further appendix 2).

Teacher marking workload was measured at baseline and post-intervention. Workload was defined as the overall number of minutes spent marking Year 5 mathematics; time spent planning was not included.[11] Teachers were asked about the overall time spent marking (in and outside of lessons) and time spent marking outside of class only. All secondary outcomes were measured using a bespoke online survey from NatCen (see Further appendix 2).

[9]KS1_MATH_OUTCOME, sourced from the NPD
[10]FSP_MAT_G11 and FSP_MAT_G12, sourced from the NPD
[11]Mathematics marking time was chosen as the focus of this measure because SDI incorporates marking into the class itself as part of the pit-stop, so it might be expected to have an impact on teacher marking time. It is important to consider the possible impact of the intervention on planning time when interpreting the impact of the intervention on marking time.

Sample size

The structure of the trial – a four-level model with randomisation at level 3 (schools) and level 4 (the three regions Lincolnshire, Derbyshire, and Oldham for stratification) – was accounted for in the sample size calculations.[12]

The calculations also assumed:

- 80% statistical power;
- statistical significance level of 95% for a two-tailed test;
- intra-cluster correlation coefficients (ICC) of 0.14 and 0.05 at the school and class level respectively;[13]
- 100% compliance with treatment assignment at the school and class level; and
- pre-test/post-test correlation of 0.1 and 0.5 at school and individual level respectively. No covariates were included at class level.[14]

At protocol stage, the total sample was 73 schools with 3,942 pupils. This yielded an MDES of 0.27[15] for the full sample and 0.30 for the FSM subgroup analysis (see Table 6).

Table 5: Minimum detectable effect size (MDES) at different stages

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
MDES		0.27	0.30	0.28	0.32	0.24	0.28
Pre-test/post-test correlations	Level 1 (pupil)	0.50	0.50	0.50	0.50	0.66	0.66
	Level 2 (class)	0.00	0.00	0.00	0.00	0.00	0.00
	Level 3 (school)	0.10	0.10	0.10	0.10	0.00	0.00
Intra-cluster correlations (ICCs)	Level 2 (class)	0.05	0.05	0.05	0.05	0.02	0.02
	Level 3 (school)	0.14	0.14	0.14	0.14	0.10	0.10
Alpha		0.05	0.05	0.05	0.05	0.05	0.05
Power		0.80	0.80	0.80	0.80	0.80	0.80
One-sided or two-sided?		2	2	2	2	2	2
Average cluster size	Classes per school	2	2	2	2	1.4	1.4

[12]The distribution of schools was unequal across the three regions, with 47% of schools being in the west, 38% in the south, and 15% in the north east.
[13]Using ICCs relating to Key Stage 2 Total Maths Scores for the north west, while class-level ICCs are expected to be smaller (EEF, 2015)
[14]Bloom *et al.* (2007) suggest that a relatively high proportion of the variance in test outcomes can be explained by pre-test scores if pre-test scores are used in adjusted analysis. Whilst including baseline attainment as a covariate can substantially reduce the individual-level covariance, accounting for this improves precision of our estimates.
[15]This number was calculated using arithmetic mean number of classes per school (1.6), rather than harmonic mean. Hence, it will differ from the actual number of pupils enrolled as described in the flow chart.

	Pupils per class	27	4[16]	27	4[16]	23	4
Number of schools	Intervention	37	37	37	37	36	35
	Control	36	36	36	36	36	36
	Total	73	73	73	73	72	71
Number of pupils	Intervention	1,998	296	1,598	237	1,339	1,325
	Control	1,944	288	1,555	230	1,491	1,489
	Total	3,942	584	3,153	467	2,830	2,814

At the time of randomisation, the mean number of classes per school was lower than anticipated (harmonic mean of 1.4 classes per school, rather than the anticipated 2). Assuming no further attrition or loss to follow-up, we estimated the study would be powered to detect an effect size of 0.28 standard deviations for the primary analysis and 0.32 for the FSM subgroup analysis

As noted above, one school dropped out mid-intervention. This meant that, at the time of analysis, our final sample was 72 schools and 3,238 pupils. There was further attrition due to loss at follow-up or measurement (14.2%), reducing the final analytical sample to 2,830 pupils and 72 schools. This smaller sample size was mitigated by a lower than expected level of clustering (0.10 at school level and 0.02 at class level). The final MDES for the primary analysis was 0.24 overall and 0.28 for the FSM subgroup analysis.

Randomisation

Randomisation was carried out by an analyst at NatCen who was blinded to the identity of schools. School identifiers were merged with group allocation data after randomisation, which was completed in Summer 2018. Schools were allocated to one of the two groups (intervention or control) using stratified randomisation by three regional hubs (Lincolnshire, Derbyshire, and Oldham) with a 50:50 ratio between the groups.[12] Stratification should have helped control for possible differences between regions in terms of school characteristics and programme implementation, but may have decreased the variance of the impact estimator (if this was correlated with the outcome). Randomisation was undertaken in STATA 14. The accompanying syntax can be found in Further appendix 5.

Statistical analysis

All the following analyses were conducted in STATA 14. An example of the syntax used for effect size calculation can be found in Further appendix 6.

Primary analysis

The primary analysis estimated the intervention's impact on Year 5 pupils' mathematics attainment using an ITT approach. *PTM10* was administered between May and June 2019. Following EEF guidance (2018), the effect size was estimated using a baseline-adjusted analysis, in which the dependent variable was the raw score of *PTM10*. This was estimated using a multilevel linear regression model. The model contained a binary indicator for treatment allocation, a

[16]Proportion of FSM pupils at protocol and randomisation assumed to be the national average of 14.4% as in Department for Education 'Schools, Pupils and their Characteristics: January 2018 – National Tables'. Totals rounded to nearest whole number.

categorical variable for KS1 attainment in mathematics[17], a categorical variable for regional hub included as a fixed effect, and random intercepts for class and school variance. The model took the following form:

$$Maths\ Attainment_{ijk} = \beta_0 + \beta_1 Baseline_{ijk} + \beta_2 Intervention_k + \beta_3 Regional\ hub + w_k + u_{jk} + e_{ijk}$$

Pupils (*i*) were clustered in classes (*j*) within schools (*k*). The estimated intervention effect was estimated as β_2. The coefficient β_3 represents the regional strata at randomisation. The model also includes random intercepts, u_{jk} and w_k, to respectively account for clustering of pupils within classes and classes within schools.

Secondary analysis

The secondary outcome analyses assessed the impact of the intervention on teachers' workload and on teachers' perceptions of pupils' confidence. These secondary outcome analyses used data from the teacher survey (see Further appendix 2). A single-level (rather than multilevel) model was used as the number of teachers per school was too small to robustly estimate random effects. Table 6 outlines response rates to both pre- and post-intervention surveys.

Table 6: Survey response rates

	Pre-intervention survey				Post-intervention survey			
	Intervention schools		Control schools		Intervention schools		Control schools	
	Teachers	SLT	Teachers	SLT	Teachers	SLT	Teachers	SLT
Total (N)	56	37	62	36	56	36	62	36
Response rate	100%	100%	100%	100%	95% (N=53)	92% (N=34)	100% (N=62)	97% (N=35)

Whilst response rates to pre- and post-intervention surveys were good, the sample for our secondary data analyses was significantly lower. This is due to the fact that, in some schools, teachers changed over the course of the academic year or were not enrolled yet when baseline surveys were conducted. Therefore, the secondary analyses using baseline data only included responses where the same teacher responded to both the baseline and post-intervention survey. Table 7 below describes the analytical sample in more detail.

Table 7: Survey analytical sample

	Intervention group	Control group	Total
Total number of teachers	56	62	118
Proportion of cases used in marking time analyses (N)	50% (N=27)	47% (N=29)	47% (N=56)
Proportion of cases used in pupil confidence analyses (N)	91% (N=51)	100% (N=62)	96% (N=113)

Once the outliers were dealt with (3%),[18] cases where teachers had answered both pre- and post-intervention surveys were included in the analyses of teachers' workload outcomes (47%). As pupil confidence was only collected at endline, it did not suffer from low rates of case matching and 96% of cases were used in the analyses. In one school, there were two teachers for a single class (working part-time). To account for this in the analysis, these teachers were assigned a weight of equal magnitude. The implications of teacher turnover and non-response are discussed in the results section and results are triangulated with findings from the process evaluation.

[17]The measure for KS1 attainment in mathematics is from the NPD database (KS1_MATH_OUTCOME). It has four categories describing if pupils have scored below, towards, as expected, or exceeded the expected level of attainment for their age group.
[18]Outliers were answers that were unusually high or low compared to the rest. When answers were obviously stated in hours rather than minutes, they were converted to minutes; when it was not obvious, they were transformed as missing.

Analysis in the presence of non-compliance

Whilst ITT analysis is informative to policymakers about the effects of an offer of treatment, it is not informative about the impact of an intervention on those who actually receive it. Consequently, we conducted an analysis of non-compliance by estimating a Local Average Treatment Effect (LATE), which is the effect of the intervention on compliers in the presence of two-sided non-compliance. This was estimated using a two-stage least squares (2SLS) instrumental-variable (IV) regression, with random assignment as the instrument.

EEF guidance (EEF, 2018b) suggests that compliance can be defined either as a binary variable or as a continuous variable. Some previous trials have used a binary definition of compliance to analyse the effect of interventions in the presence of non-compliance. However, it is often difficult to establish what a meaningful binary definition should be. Classifying cases as non-compliant when they have received a large amount of a programme risks biasing the effect estimate (Gerber & Green, 2012), whilst setting the threshold too low may also introduce bias. Therefore, we used a continuous definition of compliance, following the 'Average Causal Response to Treatment Intensity' approach outlined in Angrist and Pischke (2008). For the purpose of this evaluation, compliance was defined by two concepts:

1. the **attendance** of teachers and SLT staff at training; and
2. the **fidelity** of the delivery of the SDI.

Therefore, we interpret the LATE as the estimate of the effect of the intervention on pupil attainment when schools are compliant with the intervention (i.e. they attend training and implement core aspects of the intervention). We believe that using a continuous measure is a more nuanced approach, since a binary indicator of compliance categorising those partially compliant as non-compliant would bias the LATE estimate.

According to the ToC, training is a key part of the intervention. In the intervention group, teachers were required to attend three full days of training and SLT staff were required to attend one session. Attendance at training was recorded using registers collected by the delivery partner. Control schools could not attend training nor access any intervention material.

In addition to training, it was also important to capture whether the intervention was delivered as intended (fidelity). SDI has five key elements:

1. **Whole-class input:** use of SDI pedagogical techniques to model a concept at the start of each SDI lesson;
2. **Diagnostic assessment:** five or six carefully crafted questions are completed independently by pupils;
3. **Pit-stop:** restructuring the mathematics lesson to 75 minutes, to include a 15-minute pit-stop where the teacher marks pupils' diagnostic assessment;
4. **SDI grouping:** splitting the class into two groups based on the results of the assessment, with the teacher teaching those who still need help to fully understand a mathematics concept (the 'SDI group') and the TA supporting pupils who have already grasped the fundamentals of a mathematics concept. These pupils work on further exercises;
5. **TA:** availability of a TA for all SDI classes.

The post-intervention teacher survey (see Further appendix 2) collected data on the frequency each of these elements were used in mathematics classes for intervention and control schools (always, regularly, occasionally, or not at all). At the time of writing the SAP, we only intended to measure fidelity in the intervention schools. However, when creating the survey, we were aware that some control schools had been implementing some elements of the SDI. We therefore decided to include the same fidelity questions in the teacher survey for control schools.

As indicated in the SAP, a compliance index was constructed from measures of attendance to training and intervention fidelity. Equal weighting was given to each measure, as both were integral to the intervention (as outlined in the ToC).

The attendance score summarised teachers' and SLT staff's attendance at training for each intervention school. Each Year 5 mathematics teacher should have attended three training sessions, so each teacher was given a score out of three. Scores for each teacher were averaged across the school, resulting in one summary score per school. SLT staff should have attended one training session; if they did, one additional point was added to the score. Therefore, the overall attendance score ranged from 0 to 4, with larger values indicating higher attendance (i.e. more compliance).

Conversely, attendance at training was not required for control schools, so fully compliant control schools received a score of 0. For consistency, control schools' scores were inverted so that a higher score indicated greater compliance (as with intervention schools). As no control schools attended any training and were fully compliant with their allocation, they received a score of 4.

For fidelity, we used teachers' answers to the five questions about the SDI components listed above (coded 0 to 3, corresponding to 'not at all', 'occasionally', 'regularly', 'always'). As with the attendance score, we averaged these scores for each class and then across classes within a school, resulting in one summary score per school. The overall fidelity score ranged from 0 to 15, with larger values indicating more frequent use and therefore greater fidelity. However, this meant that, if control schools were fully compliant, their score would be 0. Therefore, scores were inverted for control schools so that a higher score indicated greater compliance for both intervention and control schools.

The two scores of attendance and fidelity were then summed to provide an overall measure of compliance. For ease of interpretation, we then rescaled the overall score so that it ranged from 0 to 1. The interpretation of the coefficient is therefore the effect we would expect the programme to have on pupil attainment in fully compliant schools.

As compliance was two-sided, we estimated a LATE using an instrumental variable approach (2SLS) with random assignment as the instrument (Angrist & Imbens, 1995). The first stage equation is as follows:

$$Comply_j = \propto + \beta_1 Treat_j + \varepsilon_{ij}$$

The predicted values of compliance $\widehat{Comply_{ij}}$ were then used in the estimation of the second stage equation, as follows:

$$Y_{ij} = \propto + \beta_1 \widehat{Comply_j} + \beta_2 Baseline_{ij} + \beta_3 Regional\ hub + \omega_{ij}$$

Pupils (*i*) were clustered in classes (*j*). $Baseline_{ij}$ indicates baseline answers and $Regional\ hub$ indicates the stratification at randomisation on regional hubs as fixed effects, with ω_{ij} representing the error term. The coefficient β_1 is our estimate of the LATE. To ensure correct estimation of standard errors with clustered data, the model was estimated with cluster-robust standard errors and using ivregress in STATA. This accounts for the clustering of pupils within schools in the estimation of standard errors but cannot fully account for the clustering of pupils within classes within schools. It is unlikely that this will substantively affect the results, but it does increase the chances of a type one error. In line with EEF guidance (EEF, 2018b) the correlation between the instrument $Treat_j$ and the endogenous variable will be reported along with the F-statistic in the analysis section.

In addition, we conducted a sensitivity analysis using a binary indicator of compliance, using the mean as a compliance threshold.

The endogeneity of compliance was tested using a Wu-Hausman test. Based on the results, the compliance estimates were revised using a standard regression model.[19]

Missing data analysis

For the primary analysis of pupil outcomes, some attrition was anticipated. This was due to external factors influencing participation in endline testing such as pupils moving schools or being absent on the day of testing. Baseline data was sourced from the NPD. Some attrition was anticipated due to pupil records not being successfully linked to the trial data, or because they were missing from baseline data. For the secondary analysis, we expected greater loss due to low and potentially differential response rates between the intervention and control group for the post-intervention teachers' survey.

For both the primary and secondary analysis, it was assumed that data was missing completely at random (MCAR) and analysis was conducted on the complete case sample. As more than 5% of the outcome data was missing, we conducted

[19] Version 1.0 of this report included the instrumental variable regression analyses, which have been moved to Further Appendix 10 in Version 2.0 of this report. Note that the mixed linear regression analysis could not be reported in Version 1.0 of the report as access to the ONS Secure Research Service was adversely impacted by the disruption caused by Covid-19. The analysis could therefore not be updated within the publication timeline.

a sensitivity analysis to explore the extent and patterns of missing data for primary and secondary analyses (as outlined in the SAP; Phillips *et al.*, 2019).

The primary analysis that was conducted used a complete case sample, which may introduce bias to our estimates. A logistic regression model was estimated for the primary outcome, to explore patterns of missing data. This model identified three characteristics associated with missing primary outcome data: age of the pupil (in months), whether the pupil had ever been eligible for FSM (EVERFSM_6_P from NPD) and KS1 mathematics attainment. As a pattern for missing data was identified, and outcome data was missing for 10.6% of pupils, multiple imputation analysis was conducted.

This analysis was conducted using Multiple Imputation through Chained Equations (MICE) using the **mi** suite of commands in STATA 14 SE. This model included the primary outcome, *PTM10* scores, and a selection of other variables.[20] The first 100 imputations were not used ('burn in') to ensure that imputation had reached a stable distribution. Whilst a stable distribution may have been achieved with fewer imputations, the benefits of a greater number of imputations were perceived to outweigh the increased computational time needed to run the multiple imputation analysis. In total, 100 imputed datasets were created to ensure the Monte Carlo error of the coefficient and standard error are reproducible within a reasonable limit (van Burren, 2018).

A logistic regression model was also estimated to explore associations with missing data for the secondary outcome, though no patterns were found. This was likely owing to the lack of covariates at teacher level. No further analysis was conducted to account for missing secondary outcome data. The secondary analysis may be subject to bias; for example, teachers in intervention schools that felt the intervention has been more useful may be more likely to respond. Therefore, we may overestimate the impact of the intervention on perceptions of pupil confidence.

Subgroup analyses

We assessed whether SDI had a different impact for pupils who had ever been eligible for FSM in the past six years (EVERFSM_6_P from the NPD). This was estimated using a similar model to the primary analysis, with the addition of the FSM indicator and an interaction term combining FSM eligibility and treatment allocation. The SAP (Phillips *et al.*, 2019) indicated that, if the coefficients resulting from this interaction were statistically significant at the 95% level, separate models would be estimated and reported for each subgroup. However, following EEF's new statistical guidance (2018b), separate models for each subgroup were also estimated, regardless of the statistical significance of the interaction coefficient. This analysis was indicative only, as the trial was powered to detect an effect for the primary outcome analysis only.

Additional analyses and robustness checks

The Statistical Analysis Plan (SAP) (Phillips *et al.*, 2019) identified several additional analyses and robustness checks. It was hypothesised that pupils participating in SDI would have a different (higher or lower) variance in attainment at follow-up, as measured by Levene's test (Levene, 1960) and the Brown-Forsythe test (Brown & Forsythe, 1974); that is, lower achievers would benefit more (or less) from the intervention than higher achievers would.

In addition, several sensitivity analyses were proposed. These were:

- an unadjusted analysis with no baseline covariates;
- an adjusted model, including a wider range of prognostic covariates to control for potential imbalance at baseline; and
- a single-level OLS regression model, using Huber-White cluster-robust standard errors and using the same covariates as the primary analysis model.

[20]Full specification: *PTM10* score, age in months at endline, the Index of Multiple Deprivation Affecting Children (IDACI) from the pupil's postcode, prior attainment in mathematics at EYFSP and KS1, Ofsted rating of the school, proportion of White British pupils at the school, proportion of pupils eligible for FSM, individual pupil eligibility for FSM in the past six years (EVERFSM_6_P), pupil gender, school type, the regional MathsHubs, and intervention or control group allocation.

Furthermore, given the change to the baseline attainment covariate in the primary analysis, an additional sensitivity analysis was conducted using both KS1 and EYFSP attainment in mathematics.

Estimation of effect sizes

Hedges (2011) constructed effect size formulae for three-level[21] cluster randomised trials and these *Hedges' g* effect sizes are used in this report. The difference in adjusted means were scaled by the pooled sample variance of the post-test measures (i.e. the unadjusted variance). The analysts used the formulae outlined below, using 95% confidence intervals.

The point estimate, *g*, was calculated as the difference between adjusted group means $\bar{Y}^T_{adj}$ and $\bar{Y}^C_{adj}$, scaled by the unconditional total standard deviation within treatment groups S_{WT}, and adjusted to account for school- and class-level clustering as follows:

$$g_{WT} = J \times \left(\frac{\bar{Y}^T_{adj} - \bar{Y}^C_{adj}}{S_{WT}}\right)\sqrt{1 - \frac{2(p_U - 1)\rho_S + 2(n_U - 1)\rho_C}{N - 2}}$$

Cohen's d can be converted to *Hedges' g* by applying the correction factor, *J*:

$$J = 1 - \left(\frac{3}{4(n_T + n_C - 2) - 1}\right)$$

The standard deviation is the square root of the estimated pooled variance, S^2_{WT}, calculated as:

$$S^2_{WT} = \frac{\sum_{i=1}^{m^T}\sum_{j=1}^{p_i^T}\sum_{k=1}^{n_{ij}^T}\left(Y^T_{ijk} - \bar{Y}^T\right)^2 + \sum_{i=1}^{m^C}\sum_{j=1}^{p_i^C}\sum_{k=1}^{n_{ij}^C}\left(Y^C_{ijk} - \bar{Y}^C\right)^2}{N - 2}$$

In these formulae, the subscripts i, j and k represent pupils, classes and schools respectively.

The school intra-cluster correlation ρ_S and the class intra-cluster correlation (ICC) ρ_C are given in the next section ('Estimation of ICC'). The remaining terms are calculated as follows:

$$p_U = \frac{N^C \sum_{i=1}^{m^T}\left(\sum_{j=1}^{p_i^T} n_{ij}^T\right)^2}{NN^T} + \frac{N^T \sum_{i=1}^{m^C}\left(\sum_{j=1}^{p_i^C} n_{ij}^C\right)^2}{NN^C}$$

$$n_U = \frac{N^C \sum_{i=1}^{m^T}\sum_{j=1}^{p_i^T}\left(n_{ij}^T\right)^2}{NN^T} + \frac{N^T \sum_{i=1}^{m^C}\sum_{j=1}^{p_i^C}\left(n_{ij}^C\right)^2}{NN^C}$$

$$N = N^T + N^C = \sum_{i=1}^{m^T}\sum_{j=1}^{P_i^T} n_{ij}^T + \sum_{i=1}^{m^C}\sum_{j=1}^{P_i^C} n_{ij}^C$$

The 95% confidence intervals were calculated as follows:

$$g_{WT} - 1.96v_g \leq \delta_T \leq g_{WT} + 1.96v_g$$

---

[21] Although this analysis accounts for four levels of clustering, the highest level is controlled using fixed effects and the three levels of random effects therefore make these formulae appropriate. The equations below are adapted from Hedges (2011).

The variance of the effect size estimate, v_g, can be conservatively approximated by:

$$v_{\{g_{WT}\}} = \frac{(1 + (p_U - 1)\rho_s + (n_U - 1)\rho_C)}{\widetilde{N}} + \frac{d_{WT}^2}{2(M^T + M^C - 2) - 1}$$

Where M^T and M^C are the number of schools in the intervention and control groups respectively. Finally, $\widetilde{N}$ is given by:

$$\widetilde{N} = \frac{N^T N^C}{N^T + N^C}$$

Estimation of ICC

The ICCs were calculated directly from the primary analysis model, using the variance estimates for each level of clustering. The formulae used to calculate the ICC for schools ρ_S and classes ρ_C are as follows:

$$\rho_S = \frac{\sigma_{BS}^2}{\sigma_{BS}^2 + \sigma_{BC}^2 + \sigma_{WC}^2}$$

$$\rho_C = \frac{\sigma_{BC}^2}{\sigma_{BS}^2 + \sigma_{BC}^2 + \sigma_{WC}^2}$$

In these formulae σ_{BS}^2 represents the between-school variance, σ_{BC}^2 the between-class variance, σ_{WC}^2 the within-class variance, and σ_{WT}^2 the sum of the variance at all levels.

Longitudinal analysis

The SDI evaluation also includes an option for follow-up analysis of pupils' mathematics attainment at Key Stage 2 (KS2). If this optional analysis were to be undertaken, the model would have the same specification as the primary analysis, only changing the outcome of interest to KS2 mathematics scores from the NPD (MATMRK). This is not assessed in this evaluation report.

Implementation and Process Evaluation (IPE)

Research methods

The IPE was designed to understand how SDI was delivered, identify any barriers or facilitators to delivery, and describe the context underpinning the delivery. The overall objectives were to explore implementation fidelity in intervention schools and describe what 'business as usual' looked like in control schools. The IPE informed analysis of the primary and secondary outcomes and explained any potential unintended outcomes. The IPE domains were decided during a workshop with the developer of the intervention and EEF. They included fidelity,[22] quality,[23] reach,[24] responsiveness,[25] and programme differentiation.[26]

The IPE research questions are listed in full in the earlier section entitled 'Evaluation objectives'.

A mixed method study comprising the following components was undertaken:

---

[22]The degree to which an intervention can be carried out in a particular setting or organisation (Proctor *et al.*, 2011)
[23]How well intervention components are implemented (Proctor *et al.*, 2011)
[24]The extent to which the intended recipients of an intervention come into contact with it (Proctor *et al.*, 2011)
[25]The extent to which the intended recipients of an intervention are engaged by the activities and content that is delivered by implementation (Proctor *et al.*, 2011)
[26]The extent to which intervention activities can be distinguished from existing practice (Proctor *et al.*, 2011)

- school and teacher pre- and post-intervention surveys that all control and intervention groups completed;
- observations of the training and surgery sessions that intervention teachers and SLT staff attended;
- case studies with ten intervention schools, including in-depth interviews with teachers, SLT staff, and TAs, and focus groups with Year 5 pupils;
- in-depth interviews with five control schools; and
- one in-depth interview with the lead developer from Yorkshire and the Humber Maths Hub.

Fieldwork was carried out between January 2018 and June 2019. Further detail about each method is set out in Table 8 below.

Table 8: IPE methods overview

Research methods	Data collection methods	Data collection sources	Sampling of data source	Achieved proportion of data collection encounters	Data analysis methods	Research questions addressed	Implementation/ ToC relevance
Pre-intervention survey	Online survey	Control / intervention teachers and members of SLT	All schools	Schools: 73 SLT 73 / 73 Teachers: 118 / 118	Regression analysis	1, 2, 3, & 5	N/A
Post-intervention survey	Online survey	Control / intervention teachers and members of SLT	All schools	Schools: 72 SLT Control: 35 / 36 Intervention: 33 / 36 Teachers Control: 62 / 62 Intervention: 53 / 56	Regression analysis	1, 2, 3, & 5	Fidelity and costs
Observations of training and surgery sessions	Observation	Training sessions	All attendees of training	Three training sessions observed and one surgery session	Review of observation templates	1,2,4,5, & 6	Fidelity
Ten qualitative case studies	In-depth interviews and focus groups	Teachers, TAs, SLT, and pupils	Purposive sampling	29 in-depth interviews 10 pupil focus groups	Framework approach	1,2,4,5, & 6	Fidelity, costs, responsiveness, and programme differentiation
Five qualitative interviews	Telephone interviews	Teachers in control schools	Purposive sampling	5 telephone interviews	Framework approach		Business as usual
Developer interview	Telephone interview	Developer	N/A	1 telephone interview	Framework approach	1,2,4,5, & 6	Fidelity, cost responsiveness, and programme differentiation

Training observations

Intervention schools were clustered by geographical location into three hubs. There were three Open Classrooms (OCs), one delivered for each hub. The remaining training was delivered to the whole cohort of intervention schools. Observations of one hub's OC, training days and surgery sessions were conducted to gain an understanding of whether the sessions were effectively building the capacity of teachers and SLT staff to deliver the intervention. In advance of the observations, a template was developed to ensure similar data was collected at each session. All teachers and SLT staff being observed were given an information sheet prior to the training commencing. This detailed the purpose of the overall evaluation, as well as the observation itself. It also explained how the data would be used in the evaluation and gave training recipients the opportunity to decline being observed if they wished. All teachers and SLT staff agreed to be observed.

SLT pre-intervention survey

The pre-intervention SLT survey (conducted between March and July 2018) was completed by SLT staff from intervention and control schools. The SLT survey collected contextual information relating to SDI delivery. The response rate was 100% as schools had to complete the survey to be included in the trial.

Teacher pre-intervention survey

The pre-intervention teacher survey (conducted between March and July 2018) was completed by Year 5 teachers in the academic year 2018/19 from intervention and control schools (see Further appendix 2). The teacher survey collected data on teacher marking workload in relation to mathematics lessons, specifically the time taken to mark. The response rate was 100% as all teachers had to complete the survey to be included in the trial.

SLT post-intervention survey

The post-intervention SLT survey (conducted between June and July 2019) was completed by SLT staff from intervention and control schools. The survey collected up-to-date information about the number of continuing professional development days teachers had received and any other mathematics interventions that had been delivered to Year 5 pupils. The SLT survey for intervention schools also asked for information about SDI delivery costs and teaching cover arrangements. The response rate was 92% for intervention schools and 97% for control schools.

Teacher post-intervention survey

The post-intervention survey (conducted between June and July 2019) was completed by teachers across intervention and control schools. Questions around teacher marking workload remained the same as those in the pre-intervention teacher survey. Intervention teachers were also asked about the extent to which they remained true to the intervention fidelity and whether any adaptations were made. The control teachers were asked whether and how often they had implemented components of SDI, if any. The response rate was 95% for intervention schools and 100% for control schools.

All survey questions were developed by NatCen and signed off by EEF. Survey questions can be found in Further appendix 2.

Case studies at intervention schools

Case studies were conducted with ten intervention schools and involved in-depth interviews with a member of the SLT, Year 5 teacher, and TA, and focus groups with Year 5 pupils. The key purpose of the case studies was to gain an in-depth insight into implementation, exploring the factors that underpin successful delivery of the intervention as well as barriers to delivery.

Purposive sampling was used to select schools, to ensure there are a wide range of experiences and circumstances across the sample using the following criteria:

- number of Year 5 classes in school – ensuring the sample includes schools with varying numbers of Year 5 classes;
- experience of training – ensuring the sample includes both schools that experienced the full complement of training as well as schools that missed some training dates; and
- hub location – ensuring the sample includes schools from across the three hubs.

Topic guides were developed for each type of data collection encounter to ensure systematic coverage of key themes that addressed the IPE research questions. An example of the topic guide can be found in Further appendix 7. Topic guides were intended to be flexible and interactive, allowing issues of relevance to be covered through detailed follow-up questioning.

Face-to-face interviews were conducted with one Year 5 teacher, a TA, and a SLT member in each case study school. The interviews aimed to gather in-depth information on school context, resource requirements, expected and perceived benefits of programme participation, implementation fidelity, and whether there were any adaptations to the delivery of SDI, what these were, and key delivery challenges and successes.

Focus groups with Year 5 pupils also took place in case study schools. Teachers were responsible for recruiting four to six pupils to participate in a group discussion which explored pupils' views and experiences of the structure and content of their mathematics lessons and their attitude and confidence towards mathematics.

Interviews at control schools

A total of five telephone interviews were conducted with Year 5 teachers in control schools. These interviews explored what business as usual looked like within control schools (see Further appendix 8 for topic guides). Purposive sampling was used to select schools, ensuring that schools varied by location and number of Year 5 classes within the school.

Each interview and discussion group lasted approximately 45 minutes. With permission from participants, all interviews and focus groups were recorded. Data was collected by members of the SDI study team, as set out in the 'Project team' section.

IPE data analysis

Quantitative data analysis

Data from the teacher and school surveys on training attendance, and delivery and implementation of SDI were also examined. Survey questions are listed in Further appendix 2. For training, we examined the proportion of teachers, TAs, and SLT staff members from intervention schools who attended each training session (see Table 22). For delivery and implementation, we calculated the proportion of respondents who, having used each of the core components of SDI, always, regularly, occasionally, not at all, or provided no response (see Table 23). We also explored business-as-usual practice in control schools by examining the proportion of teachers who used mathematics programmes such as Third Space Learning and the proportion of teachers who used teaching approaches similar to the different components of SDI (see Table 24).

Qualitative data analysis

Interviews were digitally recorded and professionally transcribed. Framework,[27] a systematic approach to qualitative data management and analysis developed by NatCen, was used to chart (collate and summarise) and analyse transcribed data. Charting in Framework involves sorting and synthesising data by theme and case, thereby helping to identify and explain patterns in the data. In this project, a mixed deductive/inductive approach was employed for the charting of transcripts, with data being synthesised according to both pre-established themes as set out in the topic guides and emerging ones. Thematic between case analysis was conducted and aimed to identify and explain patterns in the data in relation to these themes. Anonymised quotes from teaching staff and pupils were selected to illustrate key points and are included throughout the IPE section.

Costs

Collection of cost data

Cost data was collected from the post-intervention survey to estimate the cost per pupil for delivering SDI. The cost per pupil per year over three years was calculated as per EEF guidance (2016).

Calculating the average cost of delivery enables comparisons to be made with other interventions based on both the average effectiveness and costs incurred. The total cost per pupil was calculated based on information provided by schools in the school post-intervention survey about the direct and indirect costs incurred.

[27]Ritchie *et al.* (2013)

School costs

As noted, the costs information was collected in the post-intervention survey. Costs that were collected and reported in monetary terms included:

- training costs – the cost of travel and subsistence for all staff to attend all of the training;
- TA costs – the cost of hiring an additional TA to implement SDI as intended (if applicable);
- printing and photocopying costs – an estimate of the cost to print and photocopy materials for the duration of the intervention.

While the intervention did require the use of a TA in every mathematics lesson, it did not specify how schools obtained this additional resource. Not all schools hired an additional TA, so TA costs were calculated based only on the schools that reported this cost. We also asked a question about time costs: SLTs from schools were asked to estimate how many days of cover were needed to allow staff to attend training. Teachers were asked how much time they spent delivering SDI during school hours, how much time they spent outside of school hours preparing for SDI lessons, and how much time they spent in training.

Developer costs

Developer costs were collected during an interview with the developer. The costs reported are the average cost per school to deliver the intervention.

Timeline

The intervention was delivered during the 2018/19 school year (September 2018 to June 2019). The pre-intervention survey was conducted prior to randomisation, between April and May 2018.[28] Randomisation took place in July 2018. The IPE fieldwork took place between March and May 2019. Research activities concluded with post-intervention pupil testing and the SLT survey in May to July 2019. Table 9 sets out the key evaluation milestones. Analysis and reporting were delayed due to changes in the process for accessing data from the NPD. During the course of the evaluation, team members Rakhee Patel and Tanya Basi moved on from their roles at NatCen.

Table 9: Timeline of key evaluation milestones

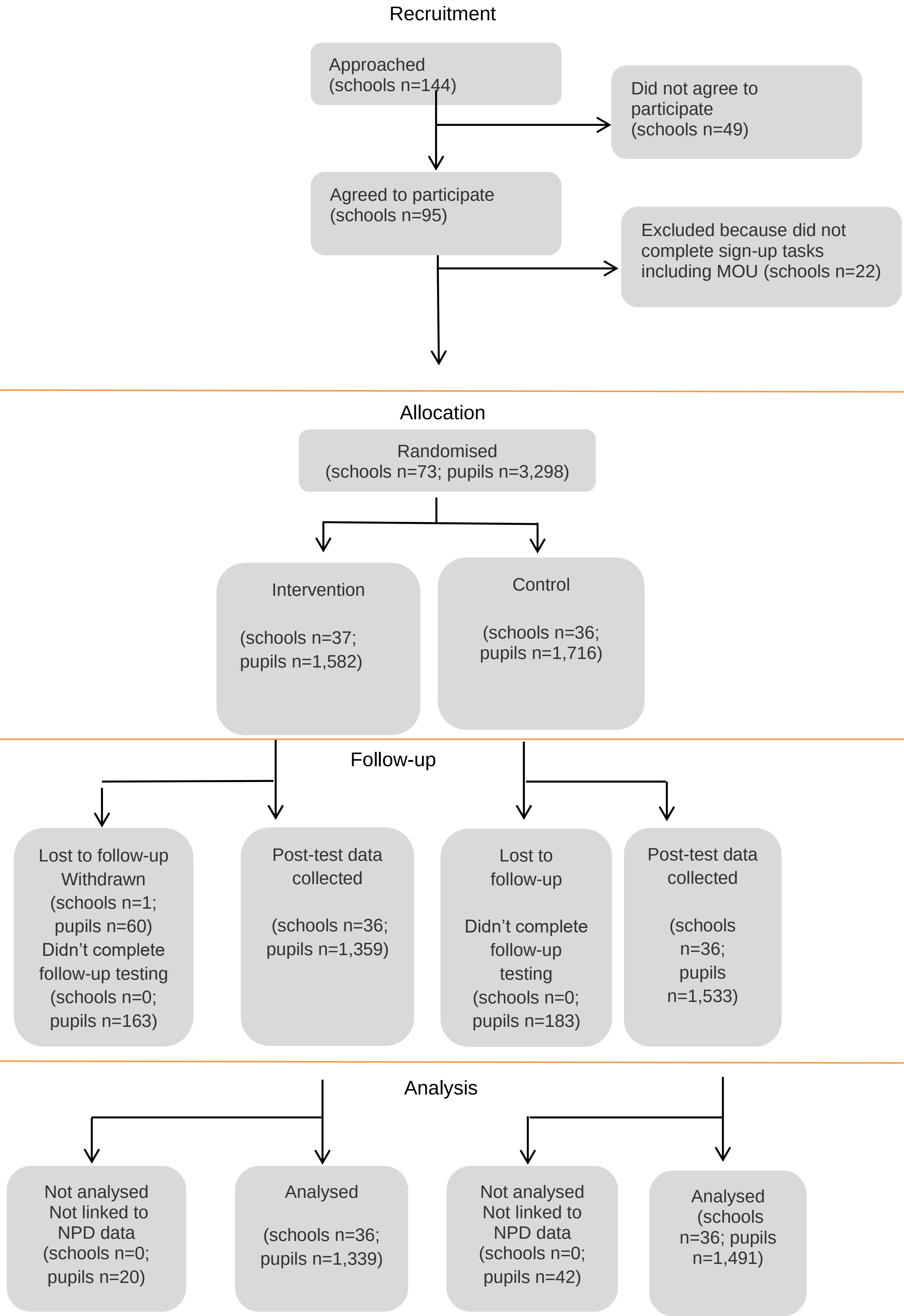
Dates	Activity	Staff responsible / leading
January 2018 to February 2018	Recruitment of schools	Rakhee Patel, Daniel Phillips, Malen Davies, Tanya Basi
April to May 2018	Pre-intervention survey	Rakhee Patel, Daniel Phillips, Tanya Basi
July 2018	Randomisation of schools	Daniel Phillips
September 2018	Training of SDI approach and SDI delivery commenced	Yorkshire and the Humber MathsHub
March to May 2019	IPE commenced	Malen Davies, Tanya Basi
May to July 2019	Post-intervention survey and intervention finished	Daniel Phillips, Shivonne Gates
June to July 2019	Testing in schools	Malen Davies, Shivonne Gates, Anysia Nguyen
September 2019 to January 2020	Analysis and reporting	Daniel Phillips, Malen Davies, Shivonne Gates, Robert Wishart and Anysia Nguyen

[28]The pre-intervention survey was conducted at the end of the school year prior to SDI implementation because delaying until the start of the following school year would have had implications for delayed randomisation, SDI training, and the start of programme implementation. However, conducting the pre-intervention survey at the end of the school year prior to SDI implementation may have contributed to the level of missing data in pupil outcomes and teacher survey datasets, as it meant there was a longer time period in which teacher changes or pupil moves between schools might occur.

Impact evaluation results

Participant flow including losses and exclusions

Figure 2: Participant flow diagram



Attrition

Table 10: Pupil level attrition from the trial (primary outcome)

		Intervention	Control	Total
Number of pupils	Randomised	1,582	1,716	3,298
	Analysed	1,339	1,491	2,830
Pupil attrition (from randomisation to analysis)	Number	243	225	468
	Percentage	15.4	13.1	14.2

Originally, 73 schools (3,298 pupils) were randomised, but one school withdrew from the trial before endline testing. Of the total pupils in the final sample, 2,892 sat the test. The remaining 346 pupils did not sit the test because they were absent due to illness or holiday, had left the school, or were pupils with SEND.[29] The final analytical sample for the primary outcome analysis was lower – 2,830 pupils in 72 schools – as some pupils could not be linked with NPD data. Consequently, the total rate of pupil attrition from randomisation to analysis was 14.2%.

Pupil and school characteristics

Randomised cluster trials, in theory, aim to achieve equal distributions of characteristics across groups at baseline. In the following tables, we checked for imbalance using pupil and school characteristics on the sample as randomised and as analysed. In Tables 12 and 13, effect sizes for difference in pupil and school characteristics are reported in the last column. An effect size of greater than 0.05 may be indicative of imbalance on these observed characteristics.

The following characteristics were selected as they are potential characteristics that could influence mathematics attainment based on our experience on previous trials (see for example Husain *et al.*, 2019). School characteristics include school type, urban-rural, Ofsted rating, and proportion of pupils eligible for FSM in a school.[30] Pupils characteristics include FSM eligibility, gender, KS1 mathematics scores, EYFSP Numbers scores, and EYFSP Shape, space and measures scores.

A single school that dropped out after randomisation objected to data processing. The descriptive analyses in Table 11 are therefore for 72 of the 73 schools randomised (3,238 pupils).[31] Table 11 indicates that the intervention and control groups were different across a range of school-level characteristics: school type, urban-rural indicator, school size as a whole and in KS2, and percentage of pupils in the school having ever received FSM. At individual level, there were small imbalances on KS1 mathematics scores and EYFSP Shape, space and measures scores. On average, pupils in the intervention group had higher scores in KS1 mathematics and EYFSP Shape, space and measures than in the control group (ES=0.09 and ES=0.12 respectively). Intervention schools had a larger proportion of pupils who had ever been eligible for FSM. Any of those imbalances could have resulted from randomisation, and they could explain differences in pupils' mathematics attainment rather than SDI. Therefore, a sensitivity analysis including all variables that were imbalanced at baseline was conducted to test whether this is the case.

---

[29]Every pupil who would usually sit in a national test like SATS was asked to sit the follow-up test. When children's needs or disabilities were too severe, teachers informed NatCen that these pupils would not sit the test.
[30]The SAP did not indicate that imbalance would be tested on urban-rural indicator, Ofsted rating, and percentage of pupils in the school having ever received FSM. However, following EEF's new report template (2019), we decided to estimate them to inform readers on whether they could apply the results to their own context. It was also observed that school size was larger in the control condition, and this information was added as it was thought to be informative.
[31]There were 38 single-form schools and 34 multi-form schools.

Table 11: Baseline characteristics of groups as randomised

Imbalance at baseline as per randomisation						
School-level (categorical)	National-level proportion (%)	Intervention group		Control group		Effect sizes
		n/N (missing)	Count (%)	n/N (missing)	Count (%)	*Hedges' g*[32]
School type						
LA maintained	N/A	16 (0)	44.44	22 (0)	61.11	-0.33
Academy	N/A	20 (0)	55.56	14 (0)	38.89	
Urban-rural						
Urban	N/A	34 (1)	94.44	31 (0)	86.11	-0.39
Rural	N/A	1 (1)	2.78	5 (0)	13.89	
Ofsted rating						
Outstanding	18	3 (10)	8.33	5 (4)	13.89	-0.01
Good	70	21 (10)	58.33	23 (4)	63.89	
Requires improvement	10	2 (10)	5.56	4 (4)	11.11	
School-level (continuous)	National-level mean	n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	*Hedges' g*
Ever FSM (%)	16	35 (1)	17.33 (12.25)	36 (0)	15.41 (11.71)	0.16
Number of pupils (KS2)	163	36 (0)	172 (85.1)	36 (0)	196 (66.6)	-0.32
Number of pupils (whole school)	282	36 (0)	319 (147.1)	36 (0)	390 (206.5)	-0.39
Pupil-level (categorical)	National-level proportion (%)	n/N (missing)	Count (%)	n/N (missing)	Count (%)	*Hedges' g*
Ever FSM						
Yes	N/A	470 (20)	30.88	531 (20)	30.94	0.00
No	N/A	1032 (20)	67.81	1165 (20)	67.89	
Gender						
Male	N/A	764 (20)	50.20	866 (20)	50.47	0.00
Female	N/A	738 (20)	48.49	830 (20)	48.37	
KS1 mathematics scores						
Below	4	54 (24)	3.55	133 (51)	7.75	0.09
Towards	18	356 (24)	23.39	365 (51)	21.27	
Expected	57	855 (24)	56.18	912 (51)	53.15	
Exceeding	19	233 (24)	15.31	255 (51)	14.86	
EYFSP Numbers						
Emerging	20	420 (51)	27.60	473 (93)	27.56	0.02
Expected	63	876 (51)	57.56	967 (93)	56.35	
Exceeding	17	175 (51)	11.50	183 (93)	10.66	

[32]The *Hedges' g* reflects the differences between the intervention and control group. For example, the effect size for the percentage of FSM pupils in schools is positive. This means that there were more FSM pupils in schools in the intervention group than the control group.

EYFSP Shape, space and measures						
Emerging	19	342 (51)	22.47	421 (93)	24.53	0.12
Expected	66	948 (51)	62.29	1066 (93)	62.12	
Exceeding	15	181 (51)	11.89	136 (93)	7.93	

Table 12, which uses only the sample as analysed, shows similar results to Table 11. This means that attrition post-randomisation did not introduce further imbalance on observed characteristics.

Table 12: Baseline characteristics of groups as analysed

Imbalance at baseline as per randomisation						
School-level (categorical)	National-level proportion (%)	Intervention group		Control group		Effect sizes
		n/N (missing)	Count (%)	n/N (missing)	Count (%)	*Hedges' g*
School type						
LA maintained	N/A	16 (0)	44.44	22 (0)	61.11	-0.33
Academy	N/A	20 (0)	55.56	14 (0)	38.89	
Urban-rural						
Urban	N/A	34 (1)	94.44	31 (0)	86.11	-0.39
Rural	N/A	1 (1)	2.78	5 (0)	13.89	
Ofsted rating						
Outstanding	18	3 (10)	8.33	5 (4)	13.89	-0.01
Good	70	21 (10)	58.33	23 (4)	63.89	
Requires improvement	10	2 (10)	5.56	4 (4)	11.11	
School-level (continuous)	National-level mean	n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	*Hedges' g*
Ever FSM (%)	16	35 (1)	17.33 (12.25)	36 (0)	15.41 (11.71)	0.15
Number of pupils (KS2)	163	36 (0)	172 (85.1)	36 (0)	196 (66.6)	-0.32
Number of pupils (whole school)	282	36 (0)	319 (147.1)	36 (0)	390 (206.5)	-0.39
Pupil-level (categorical)	National-level proportion (%)	n/N (missing)	Count (%)	n/N (missing)	Count (%)	*Hedges' g*
Ever FSM						
Yes	N/A	405 (14)	30.25	431 (2)	28.91	0.04
No	N/A	920 (14)	68.71	1058 (2)	70.96	
Gender						
Male	N/A	675 (14)	50.41	750 (2)	50.30	-0.01
Female	N/A	650 (14)	48.54	739 (2)	49.56	
KS1 mathematics scores						
Below	4	35 (0)	2.61	88 (0)	5.90	0.07
Towards	18	300 (0)	22.40	316 (0)	21.19	
Expected	57	791 (0)	59.07	845 (0)	56.67	
Exceeding	19	213 (0)	15.91	242 (0)	16.23	
EYFSP Numbers						
Emerging	20	355 (28)	26.51	375 (40)	25.15	-0.02
Expected	63	795 (28)	59.37	900 (40)	60.36	
Exceeding	17	161 (28)	12.02	176 (40)	11.80	

EYFSP Shape, space and measures						
Emerging	19	283 (28)	21.14	327 (40)	21.93	0.08
Expected	66	863 (28)	64.45	991 (40)	66.47	
Exceeding	15	165 (28)	12.32	133 (40)	8.92	

Outcomes and analysis

Primary analysis

For our primary outcome, we used the raw scores of *PTM10*, which ranged from 0 to 65. The distribution of raw scores for this test are shown in Figure 3. In both the intervention and control group, the outcome appears to have an approximately normal distribution, with no evidence of floor or ceiling effects.

Figure 3: Histogram of mathematics attainment (*PTM10* raw scores) by allocation group

The primary analysis estimated the effect size on an ITT basis. We controlled for prior attainment and the strata at randomisation, with random effects at pupil, class, and school level. The raw means and the *Hedges' g* effect size from adjusted analyses are presented in Table 13. The results show no evidence that SDI had an impact on pupils' mathematics attainment.

Table 13: Primary analysis[33]

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean [95% CI]	n (missing)	Mean [95% CI]	Total n (intervention; control)	*Hedges' g* [95% CI]	p-value
Mathematics attainment (*PTM10*)	1,339 (243)	34.5 [33.7; 35.3]	1,491 (225)	34.7 [33.9; 35.4]	2,830 (1,339; 1,491)	0.00 [-0.15; 0.16]	0.698

Possible explanations for the null findings are described in the 'Impact Process Evaluation' section and include variable fidelity in implementation across intervention schools and indications that control schools adopted elements of SDI. Alternatively, the timing of outcome testing might have been too soon to detect an effect, as many schools did not start the intervention in September as theorised in the ToC due to delays in school recruitment. There is also the possibility that the intervention had no effect. These results, as seen in the context of the whole evaluation, are further discussed in the 'Interpretation' section of the conclusion.

The full parameter of the effect size calculation can be found in the 'statistical analysis'section. The confidence intervals for *Hedges' g* effect sizes were determined using the unadjusted variance, as specified in the EEF guidance. The level of clustering observed was lower than anticipated in the power calculations, with an ICC of 0.10 at school level and 0.02 at class level.

Secondary analysis

The ToC for SDI indicates that participation is likely to improve pupils' attitude to mathematics and reduce teachers' workload. Therefore, a secondary analysis assessing the impact of SDI was suggested in the SAP. The secondary analysis looked at four outcomes relating to pupils' attitude and teachers' workload:

- teachers' time spent marking overall (in and outside class);
- teachers' time spent marking outside class only;
- teachers' perception of their pupils' confidence overall; and
- teachers' perception of their lower achievers' confidence.

The SAP (Phillips *et al.*, 2019) originally indicated that we would collect teachers' perception of pupil confidence at baseline. This was then decided against because, when teachers were asked to complete the baseline survey, they were teaching a different cohort of pupils to those who would participate in the trial. Instead, the post-intervention survey asked teachers to compare their perceptions of pupil confidence for their current Year 5 cohort (2018/19) with their previous Year 5 cohort (2017/18), based on the assumption that they were also teaching Year 5 previously. The time teachers spent marking overall and outside of class was collected both at baseline and post-intervention as this should be comparable year on year.

Due to survey non-response and teacher changes between baseline and post-intervention, the level of missing data is high for marking time outcomes (50.1%). As teacher perceptions of pupil confidence was only collected at endline, the level of missing data is much lower (4.2%).[34] As the response rates were quite low for teachers marking time and data is missing for some classes where teachers have left the school, the results of this analysis should be treated with caution. Respondents might be systematically different from non-respondents, but this hypothesis could not be explored as no covariates were collected at the teacher level.

[33]Note that the reported effect size is estimated from a multilevel regression model which includes other covariates (prior attainment and the randomisation strata). This means that the sign of the treatment coefficient used in the effect size calculation may differ from the reported difference in unadjusted means. This applies to all tables reporting means and effect sizes in this chapter.
[34]Further details on the analytical sample sizes of the secondary analyses can be found in Table 7.

The distributions of teacher marking time overall and outside of class at baseline were relatively similar for both intervention and control group (see Appendix figure 3 and Appendix figure 4). However, at follow-up, there was a statistically significant difference in marking time overall and outside of class between the two groups. In fact, teachers in the control group spent more time marking overall post-intervention (ES: -0.63 [-0.76, -0.50]). This equates to approximately an hour of marking time a week. A smaller difference was observed for marking time outside of class, with teachers in the control group still spending more time marking than those in the intervention group (ES: -0.49 [-0.63, -0.36]).

Half (50.8%) of teachers who completed the post-intervention survey did not complete the survey at baseline. This could be due to non-response, to the fact that they arrived mid-year, or because they were not assigned to a Year 5 class at the time of the baseline survey. This means that there is a lower correlation between baseline and post-intervention responses, which reduces the explanatory power of the baseline score as a covariate. Therefore, we ran sensitivity analyses that excluded baseline answers from the model to compensate for possible teacher turnover response rate. These sensitivity analyses indicated that baseline answers did not affect the outcome greatly (ES: -0.70, [-0.76; -0.63] overall; ES: -0.62 [-0.69; -0.56] outside of class).

However, it is worth noting that evidence from the IPE suggests that teachers in the intervention group initially spent more time on planning lessons as new resources had to be developed. Views on whether planning time reduced throughout the year were mixed. The 'IPE' section discusses this in more detail but, as planning time was not collected as part of the impact evaluation, it is difficult to assess quantitatively how teachers' planning workload changed.[35]

The distributions of teachers' perceptions of their pupils' confidence are available in Appendix figure 5. Results suggest that most teachers in the control group perceived their current cohort (2018/19) to be as confident as pupils from the previous cohort (2017/18). This was found for the whole cohort and for lower achievers. By contrast, a larger proportion of teachers in the intervention group indicated that their pupils were a little or a lot more confident compared with the previous cohort (both overall and lower achievers). This difference was statistically significant, with an effect size of 0.57 (CI = [0.50, 0.65]) for all pupils and 0.77 (CI = [0.69, 0.84]) for lower achievers.

Nevertheless, teachers' perceptions of their pupils' confidence is a retrospective measure. Therefore, it could suffer from confirmation bias, whereby teachers in intervention schools may answer more positively if it confirms their existing beliefs that the intervention had a positive effect. It could also suffer from recall bias as time spent marking was only recorded at endline rather than regularly. Results from this analysis should therefore be treated with this limitation in mind.

All four outcomes were regressed in a single-level cluster-robust OLS regression against the treatment allocation and the randomisation strata on an ITT. The effect sizes are presented as *Hedges' g* in Table 14 below.

[35]Overall workload was not collected as changes at that level would be too small to be detectable. Additionally, it would be impossible to unpack which aspect of workload had changed.

Table 14: Secondary analysis

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean [95% CI]	n (missing)	Mean [95% CI]	Total n (intervention; control)	*Hedges' g* [95% CI]	p-value
Teacher marking time in and outside of class	27 (29)	128.5 [100.2; 156.8]	29 (33)	190.7 [150.7; 230.6]	56 (27; 29)	-0.63 [-0.76; -0.50]	0.016
Teacher marking time outside of class	27 (29)	91.8 [63.5; 120.2]	29 (33)	130.7 [99.1; 162.3]	56 (27; 29)	-0.49 [-0.63; -0.36]	0.067
Pupils' confidence	51 (5)	2.83 [2.5; 3.2]	62 (0)	2.2 [1.9; 2.5]	113 (51; 62)	0.57 [0.50; 0.65]	0.003
Lower achievers' confidence	51 (5)	2.62 [2.3; 2.9]	62 (0)	1.8 [1.5; 2.1]	113 (51; 62)	0.77 [0.69; 0.84]	0.000

In summary, the results corroborate the conclusion of the IPE, which found that teachers reported that marking time had decreased and pupils' confidence improved. In fact, the IPE further indicates that teachers perceived that SDI built pupils' resilience, confidence, and self-esteem. It should be noted that this is the perceptions of teachers rather than the perceptions of pupils. The process evaluation section further details pupils' experience and perceptions of SDI, which this analysis does not explore.

Analysis in the presence of non-compliance

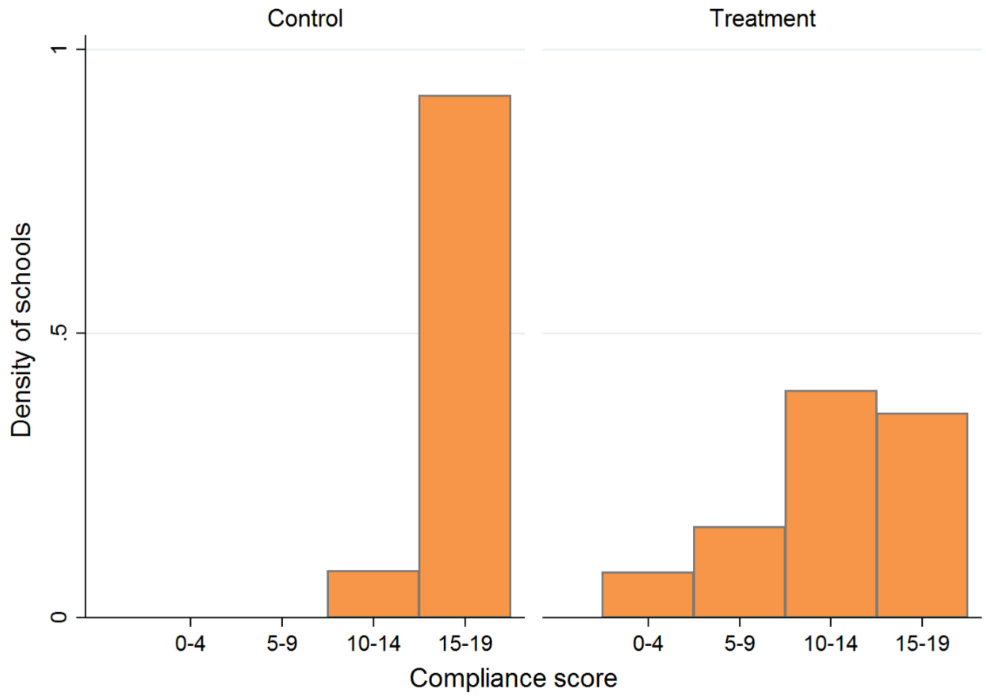
So far, the analysis has been conducted on an ITT basis. However, allocation to the intervention did not guarantee active participation in SDI, in the same way that allocation to the control group did not guarantee that some schools will not attempt to implement some aspects of SDI. The analysis accounts for two-sided non-compliance.

Figure 4 shows that, whilst the majority of control schools were compliant (i.e. did not attend any training or implement aspects of SDI), fewer than half of the intervention schools implemented SDI as intended. A full breakdown of attendance to training is provided in Table 22, whilst fidelity in intervention schools is provided in Table 23. This means that intervention schools did not necessarily implement all elements of SDI every day and some teachers or SLT staff did not attend the training sessions. The 'IPE' section on intervention delivery and fidelity provides further explanation for this partial compliance.

As SDI is composed of many elements, we decided to use a continuous measure of compliance rather than a binary measure which could be deemed more arbitrary. As detailed in the 'Methods' section, compliance was measured using a score of 0 to 19.[36] The score comprised two dimensions: attendance at training sessions and fidelity of implementation. We chose these dimensions as the ToC indicated that training sessions and programme fidelity were integral to the success of the intervention. For ease of comparison across studies, we rescaled compliance so that it ranged from 0 to 1 before estimating effect sizes.

[36]A score of 0 means that schools were not compliant at all. In the case of intervention schools, 0 means they did not attend any training or implement any aspects of SDI, whilst a score of 19 means that schools completed all training sessions and implemented all aspects of SDI as intended. In the case of control schools, full compliance (i.e. a score of 19) means that schools did not attend any training or implement any aspects of SDI.

Figure 4: Histogram of compliance measure by allocation group



Whilst the SAP (Phillips *et al.*, 2019) specified an instrumental variable approach to account for the potential endogeneity of compliance, the Wu-Hausman test indicates that compliance is actually an exogenous variable (F-statistic 0.61, p-value 0.438). Therefore, the analysis was estimated using a mixed linear regression model[37]

Table 15 Compliance analysis

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges' g (95% CI)	p-values
Mathematics attainment (PTM10)	1,339 (243)	34.5 [33.7; 35.3]	1,491 (225)	34.7 [33.9; 35.4]	2,689 (1,242; 1,502)	0.13 [-0.03; 0.29]	0.093

The results in Table 15 indicate that when accounting for non-compliance, SDI had a small impact on attainment in Mathematics, equivalent to approximately two months' additional progress in mathematics and is compatible with an effect of between zero and four additional months' progress in mathematics.

An additional sensitivity analysis used a dichotomous version of compliance (as described in the Statistical analysis section of this report). As with the primary analysis, the Wu Hausman test indicated that this should be analysed using linear regression. This alternative definition of compliance produced a lower estimate of the effect size, indicating no impact of SDI on attainment (ES: 0.03 [-0.13; 0.18]). Though it is difficult to ascertain precisely why the results differ,

[37]Version 1.0 of this report included the instrumental variable regression analyses, which have been moved to Further Appendix 10 in Version 2.0 of this report. Note that the mixed linear regression analysis could not be reported in Version 1.0 of the report as access to the ONS Secure Research Service was adversely impacted by the disruption caused by Covid-19. The analysis could therefore not be updated within the publication timeline.

we suspect dichotomising the compliance score results in lower statistical power and has less nuance relative to the continuous compliance score.

Whilst the primary analysis results indicate that the intervention had no impact, the results of the compliance analysis indicate that SDI may slightly raise attainment when implemented with greater fidelity. The null finding for the primary analysis may be explained by several factors relating to implementation. For example, the fact that schools were often unable to always have a TA available in the lesson or undertake the diagnostic test and split the class for the same day intervention (see Intervention Delivery and Fidelity, page **Error! Bookmark not defined.**). .SDI may have had a stronger effect if the programme had started when it was intended. The compliance measure did not capture the fact that only 8 percent of teachers started delivering SDI in September 2018 as required in the Theory of Change. This means that most pupils tested at follow-up in summer 2019 had not received SDI for a full academic year. Equally, outcomes may take longer to fully emerge (maturation).

Missing data analysis

As outlined in the 'Methods' section, a multiple imputation analysis was conducted using 100 imputed datasets. The analysis included a 'burn-in' of 100 to ensure the imputation model had converged to a stable distribution. Table 16 shows the results of the missing data analysis. The results of the missing data analysis confirm the null results of the primary analysis. Therefore, the findings of the primary analysis were not sensitive to missing data under the assumption that the missing data was missing at random.

Table 16: Missing data analysis

Unadjusted means				Effect sizes		
Intervention		Control				
n (missing)	Mean [95% CI]	n (missing)	Mean [95% CI]	Total	*Hedges' g* [95% CI]	p-value
1,522 (60)[38]	33.8 [33.0; 34.6]	1,716 (0)	33.4 [32.7; 34.2]	3,238 (1,522; 1,716)	0.00 [-0.23; 0.24]	0.936

Subgroup analyses

The SAP (Phillips *et al.*, 2019) outlined one subgroup analysis to determine if SDI had a differential effect on mathematics attainment for pupils who have ever been eligible for FSM (EVERFSM_6_P from NPD). This analysis used an interaction term to assess the impact of SDI on the pupils who have ever been eligible for FSM. However, the interaction coefficient was not statistically significant at the 0.05 level. Following EEF's new analytical guidance (2018), separate models were estimated for each subgroup (see Table 17). The results show a small negative association for participating in SDI whilst being eligible for FSM. For both groups of pupils, the effect sizes are very small.

Table 17: Subgroup analysis - separate models

	Unadjusted means				**Effect sizes**		
	Intervention		Control				
	n (missing)	Mean [95% CI]	n (missing)	Mean [95% CI]	Total	*Hedges' g* [95% CI]	p-value
Pupils ever eligible for FSM	405 (1,177)	29.7 [28.3; 31.1]	431 (1,285)	29.1 [27.8; 30.5]	836 (405; 431)	-0.02 [-0.22; 0.17]	0.578
Pupils never eligible for FSM	920 (662)	36.6 [35.6; 37.5]	1,058 (658)	37.0 [36.1; 37.8]	1,978 (920; 1,058)	0.01 [-0.15; 0.17]	0.544

The level of clustering observed was lower than anticipated in the power calculations, with an ICC of 0.06 at school level and 0.02 at class level for the sample that has ever been eligible for FSM.

[38]The multiple imputation analysis is missing 60 pupils due to one school dropping out after randomisation but before follow-up. Once a school withdraws we cannot process their data anymore. Therefore, we could not use any data to impute data for the 60 pupils from that school.

Additional analyses and robustness checks

One additional analysis and four robustness checks of the primary analysis were suggested in the SAP (Phillips *et al.*, 2019). The ToC explains that SDI aims to reduce the achievement gap between lower and higher achieving pupils.[39] Therefore, as SDI could affect the mathematics attainment gap in the intervention group, the additional analysis hypothesised that pupils participating in SDI will have different (higher or lower) variance in attainment at follow-up. As pupils who have ever been eligible for FSM have, on average, lower attainment than their peers there may be some overlap between this analysis and the subgroup analysis.

Table 18 shows the proportions of pupils ever eligible for free school meals in the last six years, by their classification as high or low attainers and allocation status. The results indicate that the classifications are broadly consistent across trial arms.[40]

Table 18 Ever FSM Status by prior attainment and random allocation

	Intervention		Control	
	Ever FSM	Not Ever FSM	Ever FSM	Not Ever FSM
High Achiever (%)	34	25	42	24
Low Achiever (%)	66	75	58	76
Base (n)	461	1,017	514	1,135

Base: As randomised sample, excluding cases with missing values for Ever FSM status or prior attainment

To test whether SDI closed the gap between high and lower performing pupils, we computed the Levene's and Brown-Forsythe's[41] robust test statistics to examine equality of variances between the intervention and control groups (see Table 19).

Table 19: Additional analysis: variance in attainment at follow-up

<table>
<tr><th colspan="4">Unadjusted means</th><th colspan="3" rowspan="2">Effect sizes</th></tr>
<tr><th colspan="2">Intervention</th><th colspan="2">Control</th></tr>
<tr><th>n (missing)</th><th>Mean [95% CI]</th><th>n (missing)</th><th>Mean [95% CI]</th><th>Total</th><th>F-statistic</th><th>p-value</th></tr>
<tr><td rowspan="2">1,359 (223)</td><td>34.4</td><td rowspan="2">1,533 (183)</td><td>34.4</td><td rowspan="2">2,892 (1,359; 1,533)</td><td>Levene: 0.029</td><td>0.864</td></tr>
<tr><td>[33.6; 35.2]</td><td>[33.6; 35.1]</td><td>Brown-Forsythe: 0.042</td><td>0.838</td></tr>
</table>

This analysis found no difference in the variance in the outcome between groups, indicating that there is no evidence that SDI had an impact on any attainment gap between higher and lower achieving pupils.

[39] We defined lower achievers as pupils whose KS1 mathematics score (KS1_MATH_OUTCOME) fell into the categories 'below' or 'towards' the expected marks, as defined in the NPD. Pupils classed as higher achievers are those whose marks were in the categories 'expected' or 'exceeded' in the same test.
[40] Note that Table 18 was added in version 2.0 of the report. It was not possible to add it to the earlier v 1.0 of the report as access to the ONS Secure Research Service was adversely impacted by the disruption caused by Covid-19.
[41] Brown and Forsythe's test replaces the mean in Levene's formula with the median.

As outlined in the SAP (Phillips *et al.*, 2019), we also conducted the following three robustness checks of the primary analysis plus a sensitivity analysis including imbalanced covariates:

1. an unadjusted analysis that does not include baseline covariates;
2. an adjusted model that includes a wider range of prognostic covariates: FSM eligibility, gender, school type;
3. a single-level OLS regression model, using Huber-White cluster-robust standard errors with the same variables included as in the primary model;
4. a sensitivity analysis including the variables that were imbalanced at baseline in the primary model.[42]

The results of these four analyses are presented in Table 19. The analysis was conducted using complete cases and as a consequence of missing data on some of the additional covariates, a smaller number of pupils were included in the analysis.

Table 20: Robustness checks and sensitivity analysis

	Unadjusted means				Effect sizes		
	Intervention		Control				
	n (missing)	Mean [95% CI]	n (missing)	Mean [95% CI]	Total	*Hedges' g* [95% CI]	p-value
Unadjusted analysis	1,357 (225)	34.4 [33.6; 35.2]	1,529 (187)	34.4 [33.7; 35.2]	2,886 (1,357; 1,529)	0.00 [-0.16; 0.15]	0.917
Adjusted analysis	1,325 (257)	34.5 [33.6; 35.3]	1,489 (227)	34.7 [33.9; 35.5]	2,814 (1,325; 1,489)	0.02 [-0.13; 0.18]	0.674
OLS with cluster-robust standard errors	1,339 (243)	34.5 [33.7; 35.3]	1,491 (225)	34.7 [33.9; 35.4]	2,830 (1,339; 1,491)	0.01 [-0.06; 0.09]	0.698
Baseline imbalance	1,275 (307)	34.5 [33.7; 33.3]	1,451 (265)	34.8 [34.0; 35.6]	2,726 (1,275; 1,451)	0.03 [-0.13; 0.18]	0.593

Overall, the sensitivity analyses were consistent with the primary analysis: there is no evidence that implementing SDI had an impact on pupils' mathematics attainment.

[42]These included school type, urban-rural setting, percentage of pupils eligible for FSM in schools, KS1 scores, and EYFSP Shape, space and measures scores.

Implementation and Process Evaluation results

This section synthesises the findings from the IPE to respond to the study research questions. It brings together the perspectives of school staff and pupils who participated in the evaluation. It also describes usual practice in control schools. The section concludes with a discussion of how well the findings support the ToC developed at the start of the evaluation.

Intervention delivery and fidelity

This section covers intervention delivery and fidelity by discussing:

- training and support;
- planning for the intervention;
- intervention delivery.

Training and support

The ToC stated that all teachers must attend all three training sessions and an SLT member must attend the first session. It assumed that, after participating in the training, teachers would know how to deliver the component parts of SDI and provide good quality modelling (teaching). It was compulsory for teachers to attend all three training days and optional for them to attend the surgery sessions. For SLT members it was compulsory that they attend training day 1 but attendance at any further training was optional. It was not compulsory for TAs to attend any training. It was the original intention for the training to be delivered in the Summer term, in order for teachers to begin delivering SDI as soon as the new term began. However, due to recruitment challenges outlined in the introduction, the training had to be delayed until September. This reduced the period within which SDI could be delivered from nine to seven months.

Table 21 below provides an overview of the different training components and which staff member was required to attend. Table 22 summarises the number and proportion of staff that attended the different training components.

Table 21: Summary of required staff attendance at training

	Staff attendance		
Training day	Year 5 teacher	SLT member	TA
Training day 1 – Open Classroom (OC)	Required	Required	Optional
Training day 2 – diagnostic and SDI	Required	Optional	Optional
Training day 3 – follow-on tasks	Required	Optional	Optional
TA OC	Optional	N/A	Optional
Surgery sessions	Optional	Optional	N/A
Telephone support	Optional	Optional	N/A

Table 22: Summary of actual staff attendance at training

	Staff attendance (%)		
Training day	Year 5 teacher	SLT member	TA
Total	N=55	N=37	N=37
Training day 1 – Open Classroom (OC)	73% (N=41)	60% (N=22)	0% (N=0)
Training day 2 – diagnostic and SDI	68% (N=38)	8% (N=3)	8% (N=3)
Training day 3 – follow-on tasks	73% (N=41)	5% (N=2)	8% (N=3)
TA OC	7% (N=4)	N/A	41% (N=15)
Surgery session 1	66% (N=37)	16% (N=6)	3% (N=1)
Surgery session 2	41% (N=23)	8% (N=3)	3% (N=1)
Surgery session 3	21% (N=12)	5% (N=2)	3% (N=1)
Surgery session 4	11% (N=6)	3% (N=1)	3% (N=1)

Cumulative training attendance

To give an indication of the engagement in training by teachers, a cumulative attendance score has been created where 3 indicates that all three compulsory sessions were attended, a score of 1 or 2 indicates that at least one or two sessions were attended, and 0 denotes that no training sessions were attended. In total 55% (N=31) of teachers received a score of 3, 82% (N=45) received a score of 2, 91% (N=50) received a score of 1, and 9% (N=5) did not attend any of the compulsory sessions so received a score of 0.

For SLT staff, only one session (the OC) was compulsory. As noted in Table 22, an SLT member attended at least one training session in 60% (N=22) of schools. Of these SLT staff members, one attended one additional training day and three attended two additional training days; none of the SLT members who missed the compulsory training day attended other training days in replacement.

Training day 1: Open Classroom (OC)

The developers deliberately avoided using 'outstanding' teachers to demonstrate SDI. They felt having an exceptional teacher might detract teachers' and SLT members' focus away from the delivery of the SDI components.

> *"We didn't select particularly outstanding teachers to deliver the Open Classroom […]. If you watch an absolutely outstanding teacher, you start to fixate on all the other outstanding things that they're doing in their classrooms and you're not focusing on the model of Same Day Intervention and the key ingredients that go into that."* (Developer)

Views about the OC training were mostly positive. Teachers found it useful to observe a practical demonstration of how an SDI lesson could be taught, as opposed to reading about it or being taught a theoretical concept. They also found it useful to see a demonstration of how to manage a class through each step, use the TA for support, introduce topics, and lead different components of SDI.

> *"It's probably one of the most useful aspects of the training. You can actually see it in action and then work out how it would work in your classroom."* (Year 5 teacher)

SLT members reported that the OC clarified what SDI was trying to achieve and how it should work. It also alerted them to the practical considerations they would need to address prior to implementation, such as organising the school timetable to suit the structure of SDI.

Teachers and SLT members suggested that the OC training could be improved by having a larger class size that was closer to the typical size of a Year 5 class.

> *"It would have been maybe more helpful to have a full class of 30 […] I don't know how it would work in my classroom because there's twice as many children and there's various behavioural issues."* (Year 5 teacher)

Although the developers deliberately avoided using outstanding teachers to deliver SDI, teachers and SLT members still felt the OC was unattainable. This was in part because the teaching was of a high quality, so teachers felt unable to emulate this. SLT members also noted that the lesson was unrealistic and unlikely to occur in reality owing to factors like poor or disruptive pupil behaviour.

Attendance at the OC training for TAs was optional. Not all TAs could attend OC training because they could not take time out of the school day. Where TAs did attend (49%), teachers felt able to discuss planning and teaching with TAs, comparing the OC with their own approaches. This meant the teacher and TA were able to take a more collaborative approach to adapting lessons to meet pupils' needs.

This was consistent with the observations made by TAs during interviews, who explained that the training improved their understanding of the intervention's aims. This enabled them to change and adapt their practice to more effectively support the teacher. Examples of this included a TA who, following the training, adapted the structure of their worksheets to make marking more straightforward.

Training days 2 and 3

Teachers described several ways in which the training days were useful. First, teachers found the step-by-step approach valuable: the training broke down each SDI component, which teachers felt helped to embed the different components. Secondly, teachers found the developers to be very supportive and welcomed the assistance provided with lesson planning and timings. Thirdly, teachers enjoyed the collaborative aspect of the training. This included discussing ideas, problems, and solutions with other teachers, as well as working together to create resources that were made available online for all intervention schools to access. Finally, the training gave teachers the opportunity to network and connect with other local schools and share resources.

In addition to these useful aspects of the training, teachers who were interviewed noted several areas for improvement:

- **More time spent on how to teach the 'small steps':** teachers were already familiar with this pedagogical technique for teaching mathematics, but they would have appreciated more time learning how to integrate this with SDI. Teachers reported this would have been particularly useful for planning the diagnostic task.
- **More clarity on the time needed to plan lessons:** teachers would have liked forewarning that time allocated to planning mathematics lessons would initially increase. Although teachers did not directly state what impact this had, it could be inferred that this was disruptive or had a negative impact on their time-management and workload.
- **Receiving resources ahead of time:** teachers felt that getting resources such as textbooks and example lesson plans before the training would have been more useful than receiving resources part-way through the training.
- **Observing how to teach a lesson that built on pupils' prior learning:** while teachers found observing the OC useful, the lesson they observed taught a new concept to pupils. Instead, teachers would have preferred to observe a lesson that built on pupils' previous learning. Teachers did not directly state a reason why this was preferable, but it could be because the introduction of new concepts is less common in the KS2 curriculum.

Surgery sessions, telephone support, and online resources

A total of four surgery sessions took place. Views on the usefulness of these sessions varied. When teachers found the sessions useful, this was because it was beneficial to talk with other teachers about their experiences of implementation and discuss any difficulties they were having. Teachers also reported that the sessions clarified areas of confusion around implementation, such as how to deal with pupils who were in the SDI group but were not working to KS2 level.

After the surgery sessions, teachers reported making adaptations to the intervention to suit the needs of their class. Examples of this included a teacher switching to the use of one diagnostic task and simplifying it for pupils with SEND to understand, as opposed to having two tasks for every lesson. This was found to be more inclusive for pupils and less time-consuming in terms of planning.

Teachers who were less positive about surgery sessions would have preferred the sessions to have a different focus. However, views on what the role and focus of the surgery sessions should have been varied. Some teachers expressed a preference for the sessions to focus on SDI delivery techniques (including how to support pupils with misconceptions), with less time spent on developing resources. Others would have preferred the sessions to have spent longer focusing on resource development.

Surgery sessions took place at the developer site.[43] This meant that teachers who were located further away from the developer site could not attend all the sessions. The distance would have required teachers to take a whole day out of teaching. After attending the first session, some teachers felt that it was a better use of their time to stay in school with their class.

The developer also introduced Padlet, an online portal that all intervention teachers had access to. Padlet was designed to provide ongoing support for teachers, outside of the surgery sessions. Padlet included two types of resources, those created by the developer and those designed during surgery sessions. The developer created a wide range of resources including exemplar lessons as well as a full year's scheme of work, broken down week by week, which included materials to help shape the diagnostic task as well as the Silver and Gold tasks. In addition, any resources developed during surgery sessions were uploaded to Padlet. This meant that teachers unable to attend sessions could access materials developed. Teachers were also able to ask the developer questions, and the questions and responses were open for all teachers to see. Teachers found the resources provided useful and made use of them through the intervention period. Teachers reported an initial delay in receiving any resources and suggested the process could have been improved if resources had been uploaded immediately after the training.

There was limited evidence of teachers accessing the telephone support provided. The developer explained that the bulk of the telephone support provided happened in-between the training sessions, where teachers required reassurance that they would be able to deliver the intervention as intended. It was noted by the developer that some schools were provided with support via email when needed.

Intervention delivery

This section comments on the practical delivery of the intervention components. It explores the extent to which schools maintained fidelity to the model (including dosage), as well as barriers to delivery.

Despite the commitment of intervention schools to deliver SDI as intended, a range of issues affected delivery. The post-intervention teacher survey found that dosage varied across schools, from full delivery as intended to schools dropping out of the intervention altogether. Out of 56 teachers at intervention schools, 53 completed the post-intervention survey. The results are summarised in Table 22 below.

[43] 'Twilight sessions' (listed in the original ToC) were renamed as 'surgery sessions' by the developer over the course of SDI development and implementation.

Table 23: SDI elements used by intervention schools

(Total n=53) SDI component	Always	Regularly	Occasionally	Not at all
Modelling	40% (n=21)	49% (n=26)	6% (n=3)	6% (n=3)
Included pit-stop	30% (n=16)	43% (n=23)	21% (n=11)	6% (n=3)
Marked during pit-stop	25% (n=13)	45% (n=24)	19% (n=10)	11% (n=6)
TA support in mathematics lessons	26% (n=14)	49% (n=26)	15% (n=8)	9% (n=5)
Used diagnostic task to group pupils	30% (n=16)	43% (n=23)	19% (n=10)	8% (n=4)

Input/modelling

A key part of SDI is the input, where the teacher is expected to model a mathematical concept to the whole class. According to the post-intervention teacher survey, the majority of teachers who completed the survey (n= 53) used the modelling aspect of SDI: 40% (n= 21) reported always using the technique and 49% (n=26) did so regularly. However, 12% (n= 6) either occasionally or never used the modelling component.

Teachers liked this particular part of the lesson and found the 'ping-pong' approach useful to embed learning. During the 'ping-pong', the teacher demonstrated how to solve a problem and then asked pupils to do the same. However, teachers found the suggested timing of this component was too short. Teachers reported needing longer with pupils to ensure they fully understood the concept in focus. There were examples of teachers adapting the intervention and lengthening this part of the lesson. This meant SDI was delivered over two mathematics lessons, reducing intervention fidelity.

> *"I don't end up having a pit-stop every day because of the way some of the lessons have been running."* (Year 5 teacher)

Pit-stop

A key component of SDI is the 15-minute pit-stop. This involves a pause in the lesson for teachers to mark the diagnostic task. Including the pit-stop elongates mathematics lessons from 60 minutes to 75 minutes. Developers suggested that pupils should attend assembly during the pit-stop, giving the teacher more than 15 minutes to mark all the pupils' work while pupils were out of the classroom.

The post-intervention teacher survey found that the majority of teachers who completed the survey restructured their lessons to include the pit-stop, with just under a third (30%, n=16) always including the pit-stop and almost half (43%, n=23) including it regularly. In contrast, a fifth of teachers (21%, n=11) reported only including the pit-stop occasionally while 6% (n=3) did not include it in any lessons.

The pit-stop was also not always used for its intended purpose (to mark the diagnostic task). While the majority of teachers always (25%, n=13) or regularly (45%, n=24) used it to mark the diagnostic task, a fifth only used the pit-stop to mark occasionally (19%, n=10). There were also teachers who reported never using the pit-stop to mark work (11%, n=6).

Case study schools' experiences of the pit-stop varied. When teachers were able to include the pit-stop, they reported that it worked well for marking the diagnostic task because marking did not take the full 15 minutes. However, there were challenges when it came to elongating mathematics lessons to include the pit-stop. Those who found the experience simple tended to already have longer lessons (for example, 90 minutes). Other schools found it challenging to build in the pit-stop for the following reasons:

- schools were not able to have assemblies at the same time every day, which meant mathematics lessons could not be scheduled to happen at the same time every day; and
- larger schools with two or more Year 5 classes usually had two assemblies, and therefore found it difficult to schedule an assembly that suited both classes.

Schools that faced these challenges trialled mathematics lessons at different points of the day, to decide when it would fit best. In these instances, teachers either relied upon the TA to support pupils in other activities (such as music, English skills or taking exercise) or mathematics classes were built around break or lunch times.

Where schools had not been able to change the length of mathematics lessons to include a pit-stop, they adapted the intervention. One example of this was asking pupils to mark their own work.

Diagnostic task and SDI group

The diagnostic task and the subsequent grouping of pupils based on its results was perhaps the core component of SDI. This was the point in the lesson where pupils demonstrated their understanding of the concept being taught (through the diagnostic task) and, following this, received additional support from the teacher if they had struggled (through the SDI group).

As with other components, the post-intervention teacher survey found that most schools complied with intervention fidelity and split pupils up into groups following the results of the diagnostic task. Just under a third (30%, n=16) always undertook this part of the intervention, while two-fifths regularly included this step (43%, n=23). In contrast, a fifth of schools (19%, n=10) reported that they only did this occasionally while 8% (n=4) never used this approach.

In some cases, teachers initially found planning the diagnostic task time-consuming. There were also difficulties developing a task that stretched all pupils in some way. The surgery sessions and ongoing support via telephone from the developer helped teachers overcome these difficulties.

Teachers identified two benefits of the diagnostic task. In some classes, where the range of abilities happened to be small, teachers found that all pupils had the potential to be in the SDI group. Teachers noted that this helped put pupils at ease when placed in the SDI group and reduced the stigma attached to getting things wrong. Teachers were also pleased that the task confirmed issues with fluency (that is, knowing mathematical facts and methods and recalling these efficiently) which they had anticipated were present for their pupils.

> *"I always suspected fluency was an issue. Now, through running the SDI group, I know that fluency is an issue. So quite often the children will understand the mathematical concept, but they don't have the fluency."* (Year 5 teacher)

The diagnostic task also revealed gaps in pupils' mathematics knowledge that had previously been taught in Year 4.

Views on whether the SDI group was able to help close the gap in knowledge varied. Some teachers found that the SDI group reduced gaps and misconceptions much earlier and more quickly than before.

> *"We've picked up what the gaps are, we've helped children try and improve their fluency, it sounds awful but it's like there's nowhere to hide. So, whether or not the children perceive that as a good or a bad thing, as a teacher I perceive that to be a good thing because it means we can help them with any gaps."* (Year 5 teacher)

Other teachers found that the number and range of abilities in the SDI group influenced the effectiveness of what they were trying to achieve. For example, large groups made it difficult for teachers to give all pupils the attention they needed; teachers said the optimal number of pupils in an SDI group was five to six.

The delivery of the SDI group was also sometimes influenced by the number of pupils with SEND in the group or the extent of the pupil's SEND needs. Teachers who had a small number of pupils with SEND felt able to teach them alongside other pupils in all mathematics lessons. They noted that this was a positive difference compared with usual practice, which often saw pupils with SEND separated from the whole class and supervised by the TA.

> *"I'm not thinking, oh lower-ability SEN, what on earth, what are they going to do, am I going to have to do a different task? We're all on the same page, but they might just get through a few less steps than some of the other children."* (Year 5 teacher)

By contrast, teachers who had pupils with profound SEND or a larger number of pupils with SEND alongside those with lower abilities reported difficulty managing the range of needs. In these instances, teachers resorted to using two diagnostic tasks to ensure everyone's needs were being met.

TA presence

Another core component of the intervention was the presence of a TA in the classroom. It was stated in the MoU (which all schools signed when joining the trial) that all intervention classrooms must have a TA present during Year 5 mathematics lessons. TA presence was crucial for intervention fidelity, as they supervise pupils while the teacher marks during the pit-stop and support more advanced pupils in the Silver and Gold tasks while the teacher leads the SDI group.

The post-intervention teacher survey found that a quarter of schools (26%, n=14) always had a TA in mathematics lessons, while half (49%, n=26) had a TA in the lesson regularly. In contrast, some schools only had a TA present occasionally (15%, n=8), while some did not have one at all (9%, n=5).

The ease of finding a TA to cover all Year 5 mathematics lessons varied across case study schools, which influenced the regularity that a TA was present. Schools that already had a large pool of TAs or already had TAs placed in Year 5 classrooms found it straightforward to have a TA in all Year 5 mathematics lessons. Schools with fewer TAs or with three-form entry found it challenging to provide all Year 5 classes with a dedicated TA for mathematics. Some of these schools (with the financial scope to do so) recruited TAs specifically to ensure their presence in Year 5 mathematics.

Schools with fewer financial capabilities provided a TA for SDI using resources they already had. This involved removing TAs from other classes to cover Year 5 mathematics lessons. In some cases, SLT members expressed concern that this was detrimental to classes who were then without a TA for a period of time, particularly classes that had high numbers of pupils with severe learning or behavioural difficulties. In these cases, SLT members explained that Year 5 mathematics lessons were the first thing to be deprioritised if a TA were required to cover teaching staff absence elsewhere in the school. This meant that sometimes Year 5 teachers were left without a TA during mathematics lessons.

Teachers' ability to deliver SDI as intended was influenced by limited or no TA presence. Teachers who ran lessons without a TA were not able to mark the diagnostic task and group pupils according to results. Others had a TA but only for the final part of the lesson to support pupils undertaking the Silver and Gold task. This created additional planning work for teachers, who had to give TAs detailed planning notes prior to the lesson so they were able to assist pupils with Silver and Gold tasks.

Frequency of SDI classes and length of intervention

Overall, the findings suggest there was wide variation in the extent to which intervention schools implemented SDI. Findings from the post-intervention survey outlined above show that, of the 53 intervention school teachers who responded to the survey, 20% were not consistently using the SDI components. In addition, survey responses indicate that not all schools were able to deliver SDI for the whole school year. To deliver SDI as intended, teachers should have started delivery by October 2018 and completed in July 2019. Of the 53 intervention school teachers who completed the post-intervention survey, 13 began delivery in September 2018 and 23 in October 2018. All other teachers began delivering SDI later in the academic year: 12 in November 2018 and 5 in December 2018 or January 2019. With regards to completing delivery, most teachers reported finishing in July 2019 (40 out of 53 teachers). Others finished earlier in June 2019 (four teachers) or May 2019 (four teachers). Five teachers reported ending delivery even earlier: one in February 2019, two in January 2019, and two in October 2018.

However, not all teachers who reported starting delivery later than intended finished early, and not all teachers who began delivery on time finished on time. In total, 27 teachers reported beginning and ending delivery as intended. Eight started on time but completed delivery early. Finally, 12 teachers reported starting delivery late but finishing on time, and five teachers reported starting late and finishing early. One teacher who responded to the survey questions did not complete this part of the survey.

The variance in delivery times goes some way to explaining why some teachers reported not using the components of SDI in mathematics lesson. For example, of the schools that reported ending delivery early, two reported ending it in October 2018. The same two also reported never delivering any of the components of SDI. However, this is not the only explanation as there were three schools that reported compliance with delivery time but, when it came to teaching SDI components in mathematics lessons, they reported never using at least one component of SDI in their mathematics lessons. This suggests that there were challenges that prevented delivery as intended, even when schools attempted to comply by delivering SDI for the entire school year.

Key delivery challenges

Findings from the case studies and open text analysis of the post-intervention survey indicate that a range of reasons prevented teachers from delivering SDI at a high fidelity, these included:

- **The need to meet whole school learning objectives set by SLT:** some schools were required to meet specific teaching objectives that prevented the regular use of SDI. Examples of this included a teacher being expected to implement the 'mental strategy' as part of their mathematics lessons, that is practising mental mathematics calculations (a key skill required for KS2 testing). When there were objectives like this in place, teachers felt it was necessary to deviate from the SDI lesson structure.
- **Lack of TA presence:** where teachers did not have a TA available during the mathematics lesson, they could not undertake SDI as intended with pupils.
- **A belief from teachers that SDI was too prescriptive:** findings from the post-intervention survey and qualitative case studies indicate that there were some teachers who found the structure of SDI too prescriptive. It was noted that pupils had begun to get bored of the structure or the learning style was not well-suited to every pupil. In these instances, teachers decided not to use SDI in every mathematics lesson. Instead they alternated using SDI with other teaching approaches throughout the week.
- **Pupils with SEND:** where there were either large numbers pupils with SEND or a small number of pupils with profound SEND, teachers found it challenging to deliver an SDI group that was inclusive and engaged all pupils.
- **Regular teacher absence due to training:** in some schools, teachers were attending professional development activities, which took them out of the classroom regularly. In their absence, a substitute teacher taught mathematics and did not use SDI because they had not received the training.

Usual practice

In control schools, the use of specific teaching techniques and the structure of mathematics lessons were explored through qualitative interviews and the post-intervention survey.

Use of techniques

Control school teachers described using a range of teaching techniques, including White Rose Maths, the Singapore method, Maths Mastery and Explore. Their use varied from a single technique to a combination of them, sometimes also combined with other resources. These included paper-based resources, such as place value charts and number lines, and online resources, such as classroom secrets, MathsHub resources, and e-textbooks.

Of the 28 control schools that responded to the post-intervention survey, almost two-thirds (64%, n=18) reported using mathematics programmes such as Third Space Learning, Success@Arithmetic, On Track Maths, Numeracy Ninjas, 1st Class@Number. Less commonly used programmes included Hegarty Maths and IDL Maths. By contrast, just over a third (35%, n=10) of control schools reported not using any mathematics programmes. Many of the programmes and techniques used by control schools included components similar to some of those included in SDI, particularly the 'whole-class input' modelling strategies and the emphasis on formative assessment.

For intervention schools, of the 33 that completed the post-intervention SLT survey, 24% (n=8) reported using mathematics programmes in addition to SDI. The programmes differed to those used by control schools and included Mathematics Mastery, Power of Two, and Rapid Maths.

Lesson structure

Details of a typical mathematics lesson were collected through case study qualitative interviews with control school teachers. Typical elements of mathematics lessons described by these teachers included: a 10- to 15-minute recall at the beginning of the lesson; independent activity for pupils to practise key skills; and a recap activity part-way through the lesson. Teachers' reported that their primary goal was to use these components to develop their pupils' fluency, reasoning, and problem-solving abilities in mathematics.

Although teachers did not directly compare common practice in mathematics lessons to SDI, there were similarities. The components teachers described included: input, where the teacher models a question and the pupils respond, formative assessment in the lesson to measure progress; and an intervention group. However, it should be noted that, when these components were used at control schools, they were not necessarily delivered in exactly the same way as SDI. For example, one teacher described marking in the lesson but not as part of a formalised pit-stop. Nevertheless, there were clear similarities between the components of mathematics lessons at control schools and SDI components.

The post-intervention survey asked teachers at control schools whether they included elements of SDI in their mathematics lessons. Of the 62 teachers who responded to the survey, 15% (n=9) responded "yes" to this question. These teachers were then asked how often they had used different elements of SDI, with the option to choose always, regularly, occasionally, or not at all. All nine teachers reported using two or more of these elements in their mathematics lessons either regularly or occasionally. Only one teacher reported using one element always, that is modelling. The majority of teachers (67% (n=6) to 78% (n=7)) reported using each SDI element either regularly or always. The results are summarised in Table 24 below.

Table 24: SDI elements used by control schools

(Total n=9) SDI component	Always	Regularly	Occasionally	Not at all
Modelling	11% (n=1)	56% (n=5)	22% (n=2)	11% (n=1)
Included pit-stop	0% (n=0)	11% (n=1)	56% (n=5)	33% (n=3)
Marked during pit-stop	0% (n=0)	22% (n=2)	56% (n=5)	22% (n=2)
TA support in mathematics lessons	0% (n=0)	67% (n=6)	11% (n=1)	22% (n=2)
Used diagnostic task to group pupils	0% (n=0)	44% (n=4)	22% (n=2)	33% (n=3)

Some control schools clearly use teaching practices that are similar to (or the same as) those introduced by SDI. It may be that SDI was not sufficiently different to business-as-usual practice in control schools to produce a notable effect on attainment, and this is one explanation for the evaluation's finding of no effect for mathematics attainment.

Pupils' experiences and views of SDI

The following section outlines pupils' views and experiences of the intervention, which were gathered during focus groups with a selection of pupils from each case study school. Teachers were responsible for selecting pupils and were asked to include pupils across a range of abilities.

Views on the diagnostic task were influenced by the pupil's ability to understand the task. Among those who found it challenging and reported that they were regularly in the SDI group, views were split. One group disliked the component as it caused stress or frustration when they could not understand and respond to the questions given.

> *"When I got on SDI, I was really upset because I have to do questions like I'd be able to do in Year 4 and Year 3, so I don't really like doing really easy questions."* (Year 5 pupil)

Others liked the diagnostic task as it helped them to demonstrate how much of the lesson they had understood and helped them alert the teacher that they required further support when it was needed.

Pupils who reported it to be less challenging (who were likely to be higher attainers) thought of the diagnostic task in one of two ways. Some enjoyed problem-solving and the challenge the task presented, whereas others found it tedious and wanted more challenging questions.

> *"I find it really quite boring, because actually I should be challenged a bit more like we should start on Silver."* (Year 5 pupil)

Among pupils who had experience of being placed in the SDI group during mathematics lessons, initial feelings of their placement varied. Some pupils were happy that they were going to receive extra support. Others either felt embarrassed that their lack of understanding had been exposed to their peers or felt left out from others who had been able to progress to the Silver task.

> *"I don't like how other people get on to Silver and Gold and they just go on about it and it makes other people feel left out."* (Year 5 pupil)

However, initial negative feelings did tend to subside, with pupils recognising the value of having additional support from the teacher. In these instances, pupils began to find the SDI group useful, either because the concepts could be explained at a slower pace or because the concept was taught in a different way.

> *"When we're going over the question it's more understanding then because you're going over it and they're showing you how to do it properly if you don't do it right in the diagnostic."* (Year 5 pupil)

Pupils who were frequently given the Silver and then the Gold task to complete enjoyed these tasks for a range of reasons. Reasons included the opportunity to work with others to solve problems and the gradual increase in difficulty of the questions included in both tasks.

Teacher perceptions of impact on pupils

Teaching staff views

Teachers, TAs and SLT members reported that attainment levels influenced the extent to which SDI improved self-efficacy and enjoyment of mathematics.

Self-efficacy

In line with the findings from the impact evaluation, the IPE found teachers perceived that pupils' confidence in mathematics increased. The change in structure of the mathematics lesson, to include a diagnostic task and to group pupils by performance, was generally perceived to help build the resilience, confidence, and self-esteem of pupils. Teaching staff explained that pupils gained confidence when they could see themselves progressing from the SDI group onto the Silver and Gold task during a lesson.

> *"We're seeing children that their confidence is improving and also their attitude towards a challenge is improving."* (Year 5 TA)

In some circumstances, teachers reported that pupils who were typically at the lower end of the spectrum in terms of ability were engaging more than usual in the input session to ensure that they would not be placed into the SDI group. In these situations, negative feelings about being placed in the SDI group were driving pupils to engage in the teaching element of the lesson, which is perhaps an unintended consequence of the intervention.

Teachers and TAs also noted that the regularity of having a diagnostic task led lower-attaining pupils to become more willing to attempt answering questions without support from a teacher or TA.

> *"A lot of them will still say they can't do it, but they'll have a go which is what they wouldn't do historically, so that's been better."* (Year 5 teacher)

Enjoyment of mathematics

In the qualitative interviews, teaching staff reported that pupils' enjoyment of mathematics varied and was influenced largely by pupil attainment. SLT staff, teachers and TAs reported that higher attaining pupils had begun to enjoy mathematics more. Enjoyment came from the opportunity to learn independently of the teachers, as well as take on the Silver and Gold tasks. Teachers suggested that pupils found these tasks more enjoyable because they were intended to deepen pupils' learning rather than continue working through similar level questions, which is what they would have typically done during previous mathematics lessons.

> *"…The [higher attainers] have really enjoyed the challenge of going straight into some fairly complex reasoning rather than doing lots of practice over and over again."* (Year 5 teacher)

Views on whether low-attaining pupils were getting more enjoyment from mathematics due to SDI varied. One opinion was that these pupils enjoyed mathematics more; the key reason was that the structure of the lesson provided them with reassurance that, if they could not understand a concept, they would receive intensive support from the teacher during the SDI group. These teachers also reported that, in some instances, pupils were not always in the SDI group every day and that there was some variation by task. An opposing view was that lower attaining pupils did not enjoy SDI because they either could not keep up with the pace of the lesson or found being in the SDI group each lesson disheartening.

Benefits for lower attaining pupils

Teachers' views were mixed as to whether the SDI had a positive impact upon lower attaining pupils and pupils with SEND. Among those who felt it had had a positive impact on pupils, it was noted that SDI allowed all pupils to engage in the same concepts and input as other pupils. SDI was felt to be more inclusive than previous teaching, where pupils with SEND and/or lower attaining pupils might be taken out of the class to sit with a TA and learn a different concept.

Other teachers and TAs were concerned that the intervention was not working for those learning at a level lower than KS2. It was reported that pupils at this level could rarely keep up with the input. As a result, they were continually placed in the SDI group during each mathematics lesson, which teachers and SLT members reported could be damaging to pupils' self-esteem.

Perceived outcomes for teachers

Findings from the qualitative case studies and the pre- and post-intervention surveys indicate that, overall, SDI resulted in positive outcomes in relation to increased confidence in teaching high quality mathematics lessons and a reduction in marking workload. However, planning time was reported to have increased markedly at the beginning of the trial.

Teacher confidence in teaching mathematics

A reported increase in confidence in mathematics teaching skills was widespread across the sample, which is a common finding when primary teachers engage in subject-specific CPD, regardless of content. Two factors were driving this increase in confidence.

The first factor was the training provided by the developers, which taught teachers to use a different approach to planning and delivering the main input part of a mathematics lesson. Teachers and SLT members explained that the SDI approach required teachers to break down mathematical concepts into much smaller steps than previously taught. This new process helped teachers reflect and focus on the concepts being taught and gain confidence in their ability to communicate clearly the underlying principles.

> *"I do feel a lot more confident, and it's helped because you are looking at it from the small step, and then getting bigger and bigger until you get to being able to solve a really tricky problem."* (Year 5 teacher)

The second factor underpinning an increase in confidence was the length of time and care needed to plan each individual activity. SLT members and teachers explained that lessons, particularly the input component, were being planned in much more depth compared to previous teaching. Teachers were also breaking down each concept into small chunks, which helped to enhance their own confidence in the concept they were teaching.

> *"I think I'm a better teacher, without a shadow of a doubt. I think more carefully about how I plan the maths lesson and what the small steps are in understanding for those children to be able to reach age-related expectation."* (Year 5 teacher)

There were, however, some teachers who felt that SDI had a limited influence on their confidence. Some teachers reported that they were already very capable mathematicians. Others had limited buy-in to SDI. Teachers with this had initially delivered the intervention but felt that the approach was not the most effective way of teaching mathematics. As a result, these teachers were not regularly using the components, so did not feel that the intervention had much influence on either their teaching methods or their confidence in teaching mathematics.

TA confidence in teaching mathematics

Views on whether SDI had increased TA confidence varied. TAs who had been in the role for long periods of time and had extensive experience of working with pupils at KS2 reported no change in their confidence as a result of engaging in SDI. On the other hand, TAs who were new to the role or who had previously supported pupils in KS1 noted that they had become more confident in teaching mathematics. TAs explained that they felt the need to prepare more for mathematics lessons in order to provide pupils completing the Silver and Gold tasks with the right level of support. In addition to this, some TAs were asked to support teachers to mark the diagnostic task, which they found improved the speed at which they could mark work.

Teacher workload

As discussed in the impact evaluation, survey findings indicate that mathematics marking time decreased. Across the qualitative sample the view was that initially lesson planning for mathematics had increased due to the variety of resources teachers had to develop for each component of the lesson. Resources that required development included a diagnostic task and the Silver and Gold tasks for pupils who were not placed in the SDI group.

However, views on whether planning time reduced throughout the year were mixed. Where teachers reported it had reduced over time, it was attributed to teachers becoming used to planning the resources needed, which enabled them to speed up. Another reason given was the introduction of shared resources provided by the developer or produced collectively with other teachers during surgery sessions. Teachers with this view also recognised that, while planning had increased this year, once all the resources had been developed for each lesson, planning would likely reduce for subsequent years.

On the other hand, there were teachers who reported that planning workload was higher than usual and continued to stay the same throughout the year. Teachers with this view tended to have limited or no experience of planning mathematics lessons, either because they were new teachers or because others had previously planned lessons for them.

> *"It's definitely increased my workload […] I think it's because the planning takes ages, and [..] sorting out the tasks. Finding all the questions and everything and then doing the actual lesson, which is longer than it used to be."* (Year 5 teacher)

Teachers participating in the case studies all reported that they had experienced a decrease in marking as a result of SDI. As described above, this was due to being able to mark during the pit-stop.

> *"In terms of the actual work, marking side of things, it's been absolutely great […] it's just reduced [marking], I'd rather put the effort into the planning than the marking."* (Year 5 teacher)

One view among teachers was that the increase in planning and decrease in marking had overall balanced their workload out a little.

Future delivery

Teachers and SLT members were asked whether they planned to continue with the intervention. Responses varied: some schools were planning on continuing and rolling out the SDI to other year groups and planned on using the same model; others reported that they would continue with some, but not all, aspects of SDI. The components which schools planned on keeping included the small steps approach to teaching concepts during the input and the pit-stop (which had proven useful to establish who and how much of what has been taught has been understood). Elements that schools would discontinue included the presence of a TA. Some SLT members explained that this would be too costly to have a TA in every classroom.

> *"[…] Trying to make sure that we've got the resources to be able to allow a TA to work with the Silver and Golds has been tricky in Year 5 but [it] would be even trickier [to implement across the] whole school."* (SLT member)

The post-intervention teacher survey asked a series of questions about the future delivery of SDI. Of the 56 intervention teachers who responded, the majority reported that they would recommend SDI to other schools (61%, n=34). There were three main reasons for this. First, it provided instant feedback to pupils and immediately addressed pupil's misconceptions; secondly, it reduced marking; and, thirdly, it allowed pupils of all abilities to engage in the lesson fully.

Only 11% (n=6) of teachers said they would not recommend SDI to other teachers. Teachers with this view gave one of three reasons for this. First, they believed that the planning took too long; secondly, teachers thought that the structure did not give pupils a chance to deepen their learning; and, thirdly, some teachers thought that the structure was boring for pupils, making it difficult to keep them engaged.

Finally, a quarter of respondents (25%, n=14) said they did not know if they would recommend SDI. Reasons given for this centred around the concern that planning took too long. Others noted that they did not have the resources to undertake SDI correctly, such as flexibility to change the class timetable to introduce a pit-stop in the lesson and the regular presence of a TA. In these instances, teachers were not sure how effective the intervention had been.

The survey also asked intervention teachers if they intended to continue with SDI. Intention to continue was split evenly, with 48% (n=27) saying they were going to carry on using SDI and 48% (n=27) stating they were not; two teachers did not complete the question. Teachers were also asked what changes they would make to SDI so that it would work better for their school. There were three key suggested changes. First, pupils could self-mark their diagnostic task, as this would limit the need for a pit-stop; secondly, some wanted to use the structure but make it work without a TA, as this resource could not always be guaranteed; and, thirdly, some teachers wanted to use SDI but with more flexible structure and timings. For instance, some teachers wanted to use SDI during some lessons but not all. Others wanted to increase the input session for some concepts that pupils found difficult to learn.

Summary

The key inputs that were intended to equip teachers and schools to deliver SDI included attendance at training and initial and ongoing support, both offered by the developer. Teacher attendance at the training was mixed across intervention schools so did not translate to teachers delivering SDI as intended. The post-intervention teacher survey indicates that some teachers did not use any of the SDI components or were only using them occasionally. Furthermore, research with control schools indicates that some teaching practices, which were similar to (or the same as) those introduced by SDI, were regularly part of business-as-usual mathematics teaching.

The initial ToC indicates that core SDI activities were the restructuring of the mathematics lesson to 75 minutes to include a 15-minute pit-stop for marking and the presence of a TA. The case studies found that these components were perceived by teachers as being important in allowing them to implement SDI as they had been trained. Where there was variation in the extent to which these components were used, there was limited evidence of positive outcomes from teachers and pupils. The key causal mechanisms underpinning high fidelity were ensuring a TA was present to support those pupils completing the Silver and Gold tasks so that teachers could support pupils in the SDI group and all components being delivered during every mathematics lesson.

Both the post-intervention teacher survey and the case studies indicate that there were four key reasons for low fidelity. These were: a lack of buy-in from some teachers that SDI was an effective way of teaching mathematics to Year 5 pupils; the absence of teachers from classes due to other training; a lack of TA presence to enable key elements of the SDI approach; and the inability to change the school timetable to accommodate a 75-minute mathematics class.

Perhaps the key issue affecting fidelity was a lack of TA presence during SDI classes. A lack of TA presence in SDI classrooms meant that pupils could not be split after the diagnostic tasks to allow teachers to focus on supporting those in the SDI group. Furthermore, where mathematics lessons were not split between an assembly, there was no TA to occupy pupils during the 15-minute pit-stop. Limited or lack of TA presence was largely down to schools not being able to find the resource to ensure consistent attendance in SDI classes. Even among schools that were able to resource a TA, there was appetite to adapt the SDI model in the future so that a TA was not needed, with cost often driving this consideration. Overall, it appears that many intervention schools could not support a TA's participation in SDI classes. Without extra ringfenced funding for TAs or revisions to the SDI model, it may not be possible for schools to deliver SDI as intended. It is also important to note that there is perhaps a tension between TAs being central to the intervention and its delivery, but the training for TAs being optional. Perhaps if training had been mandatory for TAs, this may have led schools to more strongly support a TA's participation in SDI classes.

A short-term impact, not articulated in the original ToC, was an increase in teacher confidence to teach mathematics. Teachers reported that their increase in confidence came as a direct result of the training received from the developer, which provided new options for planning and delivering mathematics lessons. There was also some evidence that teaching SDI had increased TA confidence in mathematics, although this was only evident for TAs who were either new to the role or new to teaching KS2.

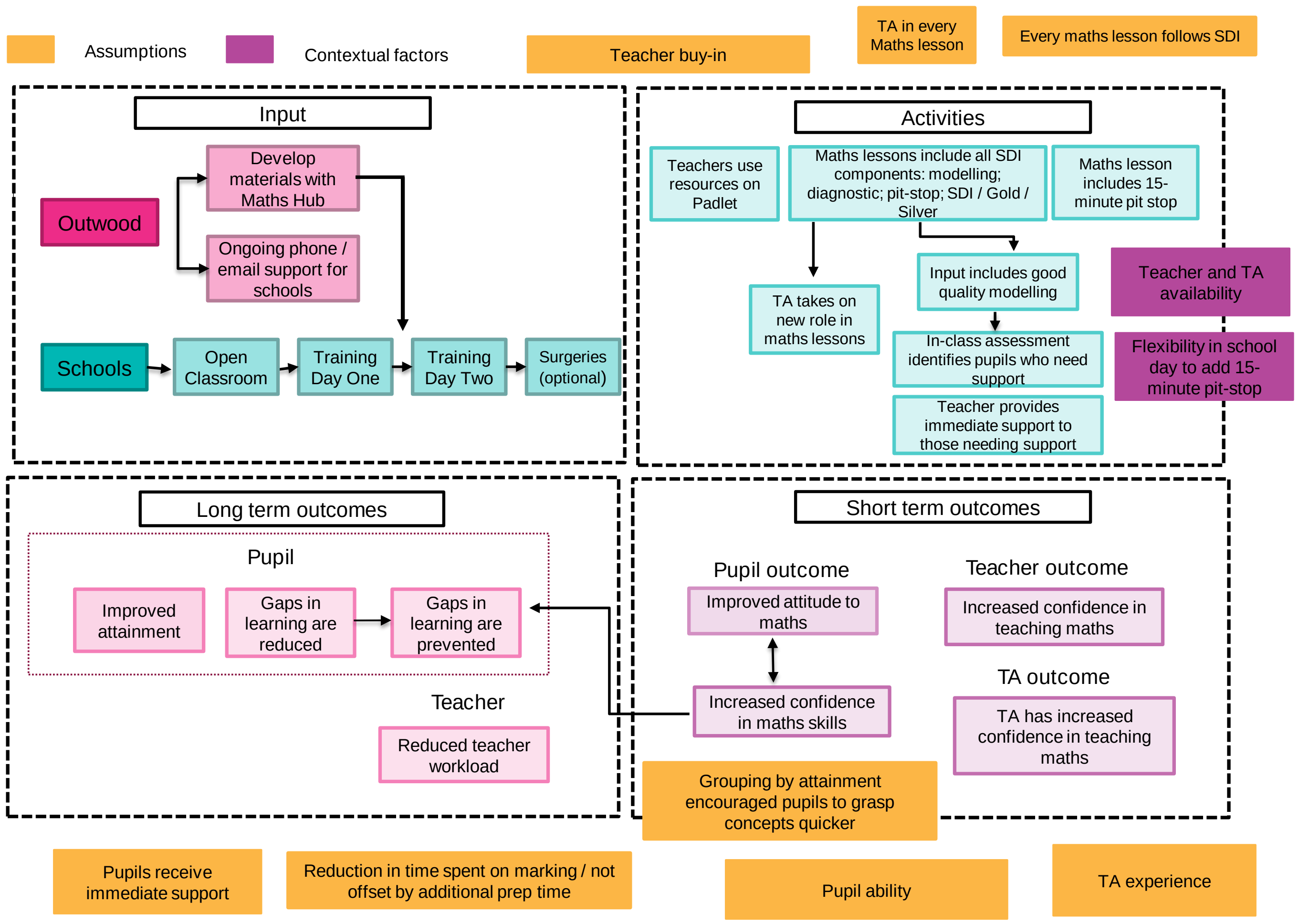
The IPE findings indicate that teachers perceived that SDI resulted in pupils having an improved attitude to mathematics and an increase in confidence in mathematics skills. This finding echoes that of the impact evaluation, which concluded that teachers' perceptions of pupil confidence in mathematics increased as a result of SDI. However, the IPE also indicates that increases in pupil confidence may have been more common for higher attaining pupils —although opinions on this were divided amongst teachers. Some teachers reported that the possibility of intensive support for those who had not understood the lesson content (i.e. lower attaining pupils) was perceived to be reassuring for pupils. In contrast, others reported that lower attaining pupils did not enjoy SDI, either because they could not keep up with the pace of the lesson or they found being in the SDI group each lesson disheartening. Pupil responses to SDI were more complex and paint a more mixed picture in relation to how receptive pupils were to the SDI model. For example, pupils had different reactions to being placed in the SDI group. Some liked the opportunity to receive more support while others were initially less positive, either because they felt left behind or embarrassed that their lack of understanding had been exposed to their peers. However, these initial feelings did tend to subside.

Another hypothesis articulated in the original ToC was that high intervention fidelity would lead to a reduction in teacher workload. The impact evaluation finds evidence of reduced teacher workload in the form of marking time. The IPE finds some evidence to support this, particularly in cases where there was high intervention fidelity. However, planning workload was also reported to have increased, especially at the initial stages of delivering the intervention. This was because teachers had to develop a wide variety of resources for different parts of the lesson, which took long periods of time to complete. Although some teachers said that time savings for marking were offset by increased preparation time, others reported that the preparation time for SDI classes reduced over the course of the school year.

Improved attitudes to mathematics were also influenced by the causal mechanism of pupils either having a positive or a negative response to being grouped after the diagnostic task. Teachers noted that lower attaining pupils were more focused and worked harder during lessons to avoid being placed in the SDI group, which could potentially increase anxiety towards mathematics.

Considering the importance of dosage and TA presence, as well as the emergence of some other evidence regarding the effect of SDI on additional outcomes, a new ToC has been developed (see Figure 5) to better articulate the different components that are critical for successful delivery of SDI and more successful outputs and impacts.

Figure 5: Revised Theory of Change



Costs

This section estimates the cost of taking part in the intervention for schools. We also estimated the time needed and the costs to the developer for delivering the training, in line with EEF cost guidance (2016).

The costs presented here are based on the costs reported by schools as being required to deliver the intervention as intended. These costs were collected through the post-intervention SLT survey (see Further Appendix 2). As described in the 'Intervention delivery and fidelity' section, there were three key requirements for schools to deliver SDI as intended. First, there were training requirements: all Year 5 teachers had to attend three whole days of training sessions and SLT staff had to attend one OC to see SDI in action. This meant that schools had to pay for teachers to travel to training. Secondly, SDI required TA support during all mathematics lessons in Year 5 classrooms, so some schools had to employ an additional TA to deliver the intervention as intended. Thirdly, schools had to adapt their mathematics lessons to follow the SDI structure, which required a new set of resources. Our estimate of the cost for schools to deliver SDI therefore includes:

- costs for staff to attend all training sessions;
- costs for schools to employ additional TAs; and
- costs for schools to print, photocopy and/or buy resources to teach SDI.

The cost estimates presented in this section are based on the number of schools, teachers, and pupils who participated in this trial. Initially, 73 schools were recruited to this trial: 36 control schools and 37 intervention schools. One school dropped out, leaving 36 intervention schools. Of these, two schools did not complete the post-intervention SLT survey. A further three schools only completed part of the SLT survey and did not provide cost data. Therefore, the cost estimates presented here are based on an average (mean) across 31 intervention schools.[44] In these 31 schools, a total of 1,268 Year 5 pupils participated in the trial and 84 school staff attended training. Our estimate of costs to schools therefore assumes the following:

- all pupils within a given Year 5 SDI class were included in the cost model;
- the number of schools delivering SDI per year and completing the costs element of the survey was 31;
- across these 31 schools:
  - the average number of pupils per school was 41;
  - the average number of staff per school that attended training was 2.33;
- the cost to attend training was the total combined cost of travel and subsistence for all staff at these 31 schools;
- the number of training days was three;
- attending training was a one-off cost every three years due to staff turnover; and
- printing resources and materials was a recurring annual cost, as resources need to be reprinted for each year group.[45]

Intended delivery was the same for every intervention school in the trial and included the three key elements described above. Table 25 summarises the costs of intended delivery of SDI based on the assumptions outlined above. The training costs described in Table 25 were calculated using the total cost of travel and subsistence reported by schools for all teachers to attend all of the sessions required. Printing and photocopying costs were calculated based on the total cost for the duration of the intervention reported by schools. Table 25 shows that the largest initial delivery cost was attending training at £814.26 per school. However, as this is only a start-up cost, the largest overall cost to schools was printing and photocopying, which was £7.32 per pupil per year over three years.

[44]We note that only four schools reported a cost above £0 in response to a survey question on the cost of any additional TA. We present the mean across the four schools that reported a cost.
[45]Although it is not an intention of SDI to increase the amount of printing for schools, teachers and SLT staff both reported that printing costs increased and that these costs would recur annually.

Table 25: Cost of delivery

Item	Type of cost	Cost per school[46]	Total cost over 3 years[47]	Total cost per pupil per year over 3 years[48]
Training: travel and subsistence	Start-up cost per school	£814.26	£814.26	£6.62
Printing / photocopying	Running cost per school	£300.32	£900.97	£7.32
Total		**£1,114.58**	**£1,715.23**	**£13.94**

Additional costs

We also collected data on the costs of employing an additional TA, as having a TA present in SDI lessons was key for intervention fidelity. However, only four schools that responded to the post-intervention survey reported incurring costs for an additional TA; all other schools reallocated TA resource, so no additional costs were incurred. For those who did employ another TA, costs reported ranged from £100 to £5000. In Table 26, we report the mean average cost of a TA per school across these four schools, along with the average costs reported in Table 25. From this we can see that the additional cost of a TA can increase the cost of delivering SDI substantially from £13.94 per pupil per year over three years to £46.87.

Table 26: Cost of delivery including TA costs

Item	Type of cost	Cost per school[49]	Total cost over 3 years[50]	Total cost per pupil per year over 3 years[51]
Training: travel and subsistence	Start-up cost per school	£814.26	£814.26	£6.62
Printing / photocopying	Running cost per school	£300.32	£900.97	£7.32
Additional TA	Running cost per school	£1,350.00	£4,050.00	£32.93
Total		**£2,464.58**	**£5,765.23**	**£46.87**

There was also the optional additional cost of having an SDI specialist attend school during the first year of implementation. This was highlighted through the interview with the developer when we collected other information about developer costs. None of the schools in the trial used a specialist, but we report the costs for information.

---

[46]EEF cost guidance (2016) requires cost per school. To determine this cost, we present a mean average. This was calculated as follows:
(total cost for all intervention schools that responded to the survey) / (total number of survey responses from intervention schools)
[47]Except where costs were considered as one-off set-up costs, total cost over three years was calculated as per EEF cost guidance (2016): *(cost per school) × 3*
[48]As per EEF cost guidance (2016), total cost per pupils over three years was calculated as follows:
(total cost over 3 years) / 3 / (total number of pupils in the trial)
[49]EEF cost guidance (2016) requires cost per school. To determine this cost, we present a mean average. This was calculated as follows:
(total cost for all intervention schools that responded to the survey) / (total number of survey responses from intervention schools)
[50]Except where costs were considered one-off set-up costs, total cost over three years was calculated as per EEF cost guidance (2016): *(cost per school) × 3*
[51]As per EEF cost guidance (2016), total cost per pupils over three years was calculated as follows:
(total cost over 3 years) / 3 / (total number of pupils in the trial)

Table 27: Cost of delivery including optional SDI specialist

Item	Type of cost	Average cost per school per year[52]	Total cost over 3 years[53]	Total cost per pupil per year over 3 years[54]
Training: travel and subsistence	Start-up cost per school	£814.26	£814.26	£6.62
Printing / photocopying	Running cost per school	£300.32	£900.97	£7.32
SDI specialist	Start-up cost per school	£1050.00	£1050.00	£8.54
Total			**£2,765.23**	**£22.48**
Additional TA	Running cost per school	£1,350.00	£4,050.00	£32.93
Total			**£6,815.23**	**£55.41**

Table 27 reports the costs of delivery noted in Table 25 plus the cost of additional support from a specialist. The estimate assumes that the specialist would visit the school for one full day per term throughout the school year (three days in total), and that the cost per day would be £350 (as stated by the developer). The additional cost of specialist support is substantial: the average cost per pupil per year over three years would increase from £13.94 to £22.48, and from £46.87 to £55.41 if schools also employed an additional TA.

Costs over time

A large part of the costs of SDI come from training (a start-up cost that recurs every three years due to staff turnover). Consequently, the per-pupil cost of delivering SDI would reduce in the second and third years of delivery. Table 27 shows that, although the average per-pupil cost is £31.43 in year 1, it subsequently only rises by an additional £11.57 per year in the following two years so that the average cost per year over three years is £43.00 per pupil.

Table 28: Cumulative costs of SDI over three years

	Year 1	Year 2	Year 3	**Average per year over 3 years**
Per pupil	£31.43[55]	£43.01[56]	£54.58[57]	**£43.00**
Per school	£1,288.77[58]	£1,763.29[59]	£2,237.81[60]	**£1,763.00**

[52]EEF cost guidance (2016) requires cost per school. For this trial, this varied across schools so here we present a mean average. This was calculated as follows:
(total cost for all intervention schools that responded to the survey) / (total number of survey responses from intervention schools)
[55]As per EEF cost guidance (2016), per pupil cost for Year 1 was calculated as follows: *(mean total cost per school that completed SLT survey) / (mean number of pupils in intervention schools that completed SLT survey)*
[56]As per EEF cost guidance (2016), per pupil cost for year 2 was calculated as follows:
(per-pupil cost for year 1) + (mean recurring cost per school) / (mean number of pupils in intervention schools that completed SLT survey)
[57]As per EEF cost guidance (2016), per pupil cost for year 3 was calculated as follows:
(per pupil cost for year 2) + (mean recurring cost per school) / (mean number of pupils in intervention schools that completed SLT survey)
[58]As per EEF cost guidance (2016), per school cost for year 1 was calculated as follows: *(mean cost per school)*
[59]As per EEF cost guidance (2016), per school cost for year 2 was calculated as follows:
(per school cost for year 1) + (mean recurring cost per school)
[60]As per EEF cost guidance (2016), per school cost for year 3 was calculated as follows:
(per school cost for year 2) + (mean recurring cost per school)

School staff time

In addition to monetary costs, school staff spent time preparing for and delivering the intervention. The survey asked teachers how many hours during the normal school week they spent delivering SDI in and out of lessons. The same questions were asked about TAs assisting their class. This data was then used to estimate time spent across the school year.[61] Estimates are shown in Table 28 and include:

- the number of days school staff (teachers and TAs) spent delivering the intervention during the school year;
- the number of days school staff (SLT staff, teachers, and TAs) spent in training to deliver the intervention;
- the number of days school staff (teachers and TAs) spent informally preparing for the intervention outside of school hours (for example, time spent on planning in addition to the normal school day) during the school year; and
- the number of days of cover teaching required to allow teachers to attend training.

The number of days school staff spent delivering the intervention in addition to the normal school day is zero as there was no requirement for teachers, TAs, or SLT staff to deliver any aspect of SDI outside of the normal school day.

The estimates are shown in number of days. The estimate of time, 5.84,[62] needed for cover assumes the mean number of days of cover needed for all Year 5 teachers at a school plus the SLT member to attend training.

Table 29: Time spent delivering the intervention

Item	Type of cost	Total time spent per school per year (days)[63]	Total time per year over 3 years (days)[64]	Total time per pupil per year over 3 years (days)[65]
Time spent delivering SDI	Ongoing cost per school	127.95	389.12	3.17
Time spent preparing outside of school hours	Ongoing cost per school	73.97	221.92	1.81
Time spent in training	Start-up cost per school	5.06	5.06	0.12
Time covering lessons	Start-up cost per school	5.84	5.84	0.14
Total			**621.94**	**5.25**

[59]As per EEF cost guidance (2016), per school cost for year 2 was calculated as follows:
(per school cost for year 1) + (mean recurring cost per school)
[60]As per EEF cost guidance (2016), per school cost for year 3 was calculated as follows:
(per school cost for year 2) + (mean recurring cost per school)
[61]This estimate was based on the assumption that one day of work is equivalent to seven hours and that there are 39 weeks in a school year. It was calculated as follows: *(total time spent in hours per week) / 7 × 39 = (total time spent in days per school year)*
[62]This mean was calculated as follows: *(total number of days of cover for all teachers at all schools that completed the SLT survey) / (total number of schools that completed the SLT survey)*
[63]EEF cost guidance (2016) requires cost per school. For this trial, this varied across schools so here we present a mean average. This was calculated as follows: *(total time spent for all intervention schools that responded to the survey) / (total number of survey responses from intervention schools)*
[64]Except where time spent was considered a one-off set-up cost, total time spent over three years was calculated as per EEF cost guidance (2016): *(time spent per school) × 3*
[65]As per EEF cost guidance (2016), total cost per pupils over three years was calculated as follows:
(total time spent over 3 years) / 3 / (total number of pupils in intervention schools that completed SLT survey)

Developer costs

The estimate of the cost for the developer to deliver the intervention shown in

includes the cost of administration time, venue hire, subsistence, and resources. This estimate assumes that:

- the cost per school for the developer to deliver training is £800 (the mean cost stated by the developer in this trial); and
- the average number of pupils per school is 41.

Table 30: Cost to the developer of delivering training

Item	Type of cost	Total cost per school	Total cost per school over 3 years	Total cost per pupil per year over 3 years
Training: admin, venue, materials, subsistence	Developer costs	£800.00	£800.00	£19.56
Total			**£800.00**	**£19.56**

Conclusion

Table 31: Key conclusions

Key conclusions
1. Pupils in Same Day Intervention schools made the equivalent of zero additional months' progress in mathematics, on average, compared to pupils in control schools. This result has a moderate to high EEF security rating.
2. Half of teachers in Same Day Intervention schools provided data on the time they spent marking pupil work, and there was evidence that they spent less time marking than teachers in control schools (approximately one hour less per week). Teachers in Same Day Intervention schools were also more likely than teachers in control schools to report that all pupils, and lower achieving pupils, were more confident compared to the previous year's pupils.
3. Exploratory analysis suggests that pupils eligible for Free School Meals (FSM) in Same Day Intervention schools made the equivalent of zero months' progress in mathematics, on average, compared to FSM pupils in control schools.
4. Fewer than half of intervention schools implemented Same Day Intervention as intended. Schools faced several challenges which made implementation difficult, such as being able to provide a teaching assistant to support delivery and adapting the school day to extend mathematics lessons.
5. Compliance analysis suggests that pupils in schools that implemented Same Day Intervention as the programme developer intended may have made, on average, two months additional progress in mathematics compared to pupils in control schools. There may be other differences in schools that implemented the programme with greater fidelity, however, making this finding less secure than the primary analysis.

Impact evaluation and IPE integration

Evidence to support the Theory of Change

In this trial, Same Day Intervention (SDI) classes were intended to replace business-as-usual mathematics classes over a period of approximately nine months, although delayed recruitment meant that the programme lasted approximately seven months. The intervention involves teachers demonstrating a given mathematics concept before pupils are given diagnostic questions. Teachers mark pupils' work and group them according to their diagnostic activity performance. The teacher provides a targeted session for lower performing pupils, with the aim of addressing common misconceptions and embedding learning. Pupils who demonstrate that they have already grasped a concept are given more challenging tasks, designed to deepen and master the topic, supported by the TA.

The programme's original Theory of Change (ToC) makes the central aim of the programme increasing pupils' mathematics attainment and providing targeted support so that all pupils attain a given level of mathematics understanding and thus prevent any achievement gap from emerging. The ToC also anticipates that SDI may improve pupils' attitude to mathematics and that the intervention may help reduce teacher workload by reducing the time spent marking pupils' work.

This evaluation did not find evidence that SDI had an impact on pupils' mathematics attainment or on closing any attainment gap. However, there were several issues with intervention fidelity. The results from the compliance analysis indicate that where the intervention was implemented with greater fidelity, pupils made, on average, two months' additional progress in mathematics, though the results are consistent with an effect of between zero and four months' additional progress. The evaluation also found an impact on teachers' perceptions of their pupils' confidence in mathematics. The IPE further found that teachers reported SDI was beneficial for mathematics self-efficacy and enjoyment of mathematics for some pupils although, while some lower attaining pupils responded positively to being put in the SDI group, others found it stressful. However, teachers' views on whether SDI was beneficial for lower attaining pupils were divided. One view was that the possibility of intensive support for those who had not understood the lesson content was reassuring. Conversely, others reported that lower attaining pupils did not enjoy SDI, either because they could not keep up with the pace of the lesson or they found being in the SDI group each lesson disheartening.

Half of the teachers provided data on the time they spent marking pupils' mathematics work and for these there was evidence of a positive impact, though this finding could be affected by non-response bias. There is also evidence from the IPE that some teachers experienced an increase in the preparation time required for mathematics classes, with only some teachers reporting that this then reduced over the course of the intervention. The IPE further found that teachers reported an increase in their own confidence in teaching mathematics – something they typically related to having received training and having learnt new options for planning and delivering mathematics lessons.

SDI training received mixed attendance from teachers (91% of treatment teachers attended at least one of three 'compulsory' sessions, but only 55% attended all three sessions), and members of schools' SLT (SDI training was attended by SLT representatives from 60% of schools). It is possible that greater attendance at training might have increased the likelihood or strength of an impact on outcomes of interest. The training was generally positively received by participants. Teachers found the step-by-step approach valuable and the developers to be supportive, and enjoyed the collaborative aspect of the training and the opportunity to build a network and share resources with other local school staff. TAs were not required to attend training, but feedback for the small numbers who did attend shows that they found the experience useful and that it helped teachers and TAs plan SDI lessons. As TAs play such a prominent role in the intervention, their attendance at training might help promote intervention implementation fidelity.

Feedback also indicated that the following improvements to training might be useful to teachers: an increased focus on how to integrate mathematics pedagogy to the SDI lesson structure and diagnostic task; greater warning that SDI might necessitate an increase in lesson planning time, at least in the short term; early receipt of SDI resources such as textbooks and example lesson plans; a demonstration of how to teach an SDI lesson that built on pupils' prior learning (in addition to examples of how to teach a new mathematics concept).

There were some issues with SDI implementation fidelity in schools. For example, not all schools were able to provide TAs for all classes. This was a fundamental obstacle to implementing the intervention as planned. It meant that SDI classes could not be implemented as intended, namely a TA looking after the pupils during a 15-minute pit-stop while the teacher marked their diagnostic tasks, and the TA and teacher each assisting separate groups of pupils with targeted help. This may have contributed to the null finding of the primary analysis.

Furthermore, some schools did not implement SDI for the anticipated period of approximately seven months. SDI classes were also not always extended so that they lasted the intended 75 minutes; some schools found this difficult to do as it involved amending the structure of the school day. Schools did not always implement the key elements of the programme such as the whole-class input or approach to modelling mathematics concepts, the diagnostic test of pupils' understanding and accompanying pit-stop to mark pupils' tests, and the division of the class into targeted groups according to the result of their diagnostic test. Inability to provide a TA for all classes or to extend classes to last 75 minutes may have contributed to the fact that some schools did not regularly implement these key elements.

Interpretation

There is little prior evidence exploring the effectiveness of Same Day Intervention. An evaluation of a pilot in schools supported by the Yorkshire and the Humber MathsHub in the north of England found some evidence that pupils' attitude towards mathematics became more positive, and mathematics attainment increased, when pre-intervention and post-intervention scores were compared (McGarry, 2017). However, this evaluation did not seek to compare an SDI intervention group with a group of schools not implementing the programme. When the broader evidence base is considered, there is support for the effectiveness of some of the key features of the SDI approach such as providing feedback, within class attainment grouping and use of TAs for small-group activities.

However, this impact evaluation finds no impact of SDI on pupils' mathematics attainment. This may result from the fact that not all schools implemented SDI as planned, or from outcome measurement being too early for an impact to emerge, or because the programme, as implemented here, is not better or distinct enough from current best-practice mathematics lessons. We find evidence that attendance at training could have been higher. We also find that the programme was not consistently implemented by schools as designed. In some schools, SDI implementation took place over a time period of less than seven months. Furthermore, some schools did not implement all key components of the SDI approach (such as regular use of a diagnostic test of pupils' understanding and the division of the class into targeted groups for further teaching) in those classes that did take place. Compliance analysis indicated that when the intervention was implemented with greater fidelity, pupils made, on average, approximately two months' additional progress in mathematics. However, this analysis could not account for other areas where implementation may not have been as intended, such as SDI implementation starting later or ending earlier than planned. As such, the intervention may be effective if it were fully implemented in schools, though this cannot be assessed in this evaluation.

Some schools found it difficult to implement SDI as per programme theory due to the requirement for a TA to be present in each SDI class. The IPE finds evidence that larger schools or those that had a pool of available TAs to draw on had fewer problems in this respect. However, some smaller schools, or those which did not have TAs before they took part in SDI, reported difficulties in resourcing this element of the programme. Without the presence of a TA, it was often difficult to implement SDI classes as intended. Another difficulty resulted from the requirement that mathematics classes last 75 minutes in order to allow for the 15-minute pit-stop and this was not always easy to schedule. The evaluation also finds that some schools started the programme later than the originally anticipated starting point of October 2019 and finished implementing it before the end of the school year. Other limiting factors on implementation fidelity included the fact that some teachers felt SDI to be too prescriptive, the need to meet wider school learning objectives, teacher absence or illness, and the difficulty of implementing the programme where there was a large proportion of pupils with SEND or a small number of pupils with profound SEND.

It may be that some of the schools included in this evaluation were drawn from Opportunity Areas or from other disadvantaged communities and required additional support to implement the programme as intended – for example, by making funding available specifically to provide TAs.

The evaluation does find an impact on teachers' perceptions of pupil confidence in mathematics – a finding supported by the IPE. It may be that the programme had enough time to produce an impact on pupil confidence but that this did not translate into an impact on mathematics attainment within the study's timeframe.

Finally, the evaluation also finds an impact on teacher marking workload, although only half of teachers provided data on the time they spent marking mathematics work. This finding is supported by the IPE, but there is also evidence that mathematics preparation time increased for some teachers.

Overall, this evaluation indicates that SDI was not effective in increasing pupils' mathematics attainment but provides some indications of impact on intermediate outcomes such as teacher marking time and teachers' perceptions of pupil confidence. Thus, the study is unable to replicate the promising finding from the pilot study by McGarry (2017). Prior evidence indicates that interventions that adopt a similar model can influence mathematics attainment, but do not always do so. We find that, in the case of SDI, some key elements of the programme were not implemented as planned and that this may have limited the intervention's potential impact on pupil mathematics attainment.

Limitations and lessons learned

Impact evaluation

The impact evaluation is subject to some methodological limitations. Due to attrition, 14.2% of randomised pupils were not included in the final analysis as they did not have the required data. Although no differential attrition was observed, the attrition may still introduce bias into the estimates (such as imbalance on unobserved characteristics). Multiple imputation analysis found results for pupil mathematics attainment consistent with the primary analysis. There was also some imbalance at baseline. Schools in the intervention group were more likely to be academies and more likely to be in urban areas, relative to the control group. However, sensitivity analysis for pupil mathematics attainment is also consistent with the primary analysis.

As a result of fewer schools being recruited than intended and the subsequent attrition, the trial is underpowered. This reduces the chances of detecting the impact of the intervention if one exists. Furthermore, mathematics attainment scores were measured relatively soon after SDI was implemented, and it may be that effects would take longer to emerge or that the intervention would need to be implemented for longer to produce an effect.

There are also relatively high levels of missing data for intermediate outcomes such as teacher marking workload and teacher perceptions of pupil confidence. These findings should therefore be interpreted with caution. The measure of teacher workload (i.e. time spent marking) does not take into account possible impact of the intervention on class preparation time. The implementation of some SDI elements in the control arm may also have introduced some contamination, as some control schools reported that they adopted some elements of SDI. However, the compliance analysis indicated that when accounting for this, the results were compatible with an impact of between zero and four months' additional progress in mathematics. With greater fidelity, we may have observed stronger impacts. It is also possible that some of the alternative interventions adopted by control schools counteracted any effect of SDI in treatment schools or that their standard mathematics classes incorporated processes and techniques similar to those used by SDI (something borne out by the interviews undertaken with staff in a sample of control schools). Furthermore, it is possible that some important determinants of SDI implementation fidelity were not accounted for in our measure of compliance.

For example, the compliance measure did not record at what point in the year the teacher started to implement SDI or stopped using the approach. Additionally, as the measure is self-reported by teachers, it is possible that this may introduce bias (e.g. desirability bias to be seen as complying with the condition they were assigned to).

Implementation and Process Evaluation (IPE)

While it was not possible for the research team to know that implementation of SDI was limited in some schools, on reflection, the purposive sampling for the IPE case studies could have selected schools based on the extent to which they were implementing all SDI components. This would have helped gain a deeper insight into why some components of SDI were not being delivered as intended.

Generalisability

SDI was originally developed and piloted in schools in Yorkshire and the Humber who were part of the Yorkshire and the Humber MathsHub Work Group. This study evaluated the effect of SDI on schools in the north of England, particularly in Yorkshire and the Humber and surrounding areas. Schools had to have at least one class of Year 5 pupils and be mainstream non-selective and non-special schools. The findings of this study may not be generalisable to non-mainstream schools or to schools outside the north of England, and specifically those in the three intervention hubs of Lincolnshire, Derbyshire, and Oldham. As noted earlier in the report, recruitment of schools to the trial was quite difficult, with many schools that initially expressed an interest in the intervention choosing not to complete all necessary sign-up tasks. It may be that the final sample of schools included the most proactive and well-resourced schools.

Future research and publications

The difficulty experienced in recruiting schools should be an important consideration for any future evaluations. Future research could explore whether SDI has an impact on teacher workload, once any additional preparation time for SDI classes is accounted for. Research could also explore whether SDI may impact on pupils' mathematics attainment if schools are provided with greater support to implement the programme as intended (such as additional funds to ensure that a TA can support all SDI classes), or explore the effectiveness of an adapted version of the intervention that does not require longer lesson length or the presence of a TA. Analysis could also be undertaken to assess if SDI could impact on attainment over a longer follow-up period, by analysing data from the NPD.

References

Angrist, J.D. and Pischke, J.-S. (2008) *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, New Jersey: Princeton University Press.

Angrist, J. and Imbens, G. (1995) 'Two-stage least squares estimation of average causal effects in models with variable treatment intensity.' *American Statistical Association,* 431–42.

Bloom, H.S., Richburg-Hayes, L. and Black, A.R. (2007) 'Using covariates to improve precision for studies that randomize schools to evaluate educational interventions.' *Educational Evaluation and Policy Analysis,* 30–59.

Brown, M. and Forsythe, A. (1974) 'Robust tests for the equality of variances.' *Journal of the American Statistical Association,* 69, 364–7.

Davies, M., Phillips, D., Nguyen, A. and Wishart, R. (2019) *Trial Evaluation Protocol – Same Day Intervention.* London: Education Endowment Foundation.

EEF (2015) *Intra-cluster Correlation Coefficients.* London: Education Endowment Foundation.

EEF (2016) *Cost Evaluation Guidance.* London: Education Endowment Foundation.

EEF (2018a) *Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.* London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit

EEF (2018b) *Statistical Analysis Guidance for EEF.* London: Education Endowment Foundation.

EEF (2019) *Impact Evaluation Report Template.* London: Education Endowment Foundation.

Gerber, A.S. and Green, D.P. (2012) *Field Experiments: Design, Analysis and Interpretation.* s.l.:W.W. Norton.

Hattie, J. and Timperley, H. (2007) 'The power of feedback.' *Review of Educational Research,* 81–112.

Hedges, L.V. (2011) 'Effect sizes in three-level cluster-randomized experiments.' *Journal of Educational and Behavioral Statistics,* 346–80.

Hodgen, J., Foster, C., Marks, R. and Brown, M. (2018) *Evidence for Review of Mathematics Teaching: Improving Mathematics in Key Stages Two and Three: Evidence Review.* London: Education Endowment Foundation.

Husain, F. *et al.* (2019) *Fit to Study.* London: Education Endowment Foundation.

Levene, H. (1960) Essays in honor of Harold Hotelling. In: *Contributions to Probability and Statistics.* Stanford: Stanford University Press, pp. 278–92.

McGarry, B. (2017) *Same Day Intervention.* Yorkshire and the Humber: MathsHubs.

NCETM (2015) *How can we meet the needs of all pupils without differentiation of lesson content? How can we record progress without levels?* Viewed 07 May 2020, <https://www.ncetm.org.uk/resources/46830>: s.n.

Phillips, D., Wishart, R. and Nguyen, A. (2019) *Statistical Analysis Plan: Same Day Intervention.* London: Education Endowment Foundation.

Proctor, E. *et al.* (2011) 'Outcomes for implementation research: Conceptual distinctions measurement challenges and research agenda. *Administration and Policy in Mental Health,* 65–76.

Ritchie, J., Lewis, J., McNaughton, N. and Ormston, R. (2013) *Qualitative Research Practice: A Guide for Social Science Students and Researchers.* s.l.:Sage Publications Ltd.

Soderstrom, N.C. and Bjork, R.A. (2013) 'Learning versus performance.' In: *Oxford Bibliographies Online: Psychology.* New York: Oxford University Press.

van Burren, S. (2018) *Flexible Imputation of Missing Data.* London: Chapman and Hall/CRC Press.

Appendix A EEF cost rating

Appendix table 1: EEF Cost rating

Cost rating	Description
£ £ £ £ £	*Very low:* less than £80 per pupil per year
£ £ £ £ £	*Low:* up to about £200 per pupil per year
£ £ £ £ £	*Moderate:* up to about £700 per pupil per year
£ £ £ £ £	*High:* up to £1,200 per pupil per year
£ £ £ £ £	*Very high:* over £1,200 per pupil per year

Appendix B Security classification of trial findings

OUTCOME: *Mathematics attainment*

Please use this template to assign a separate security rating for each primary outcome.

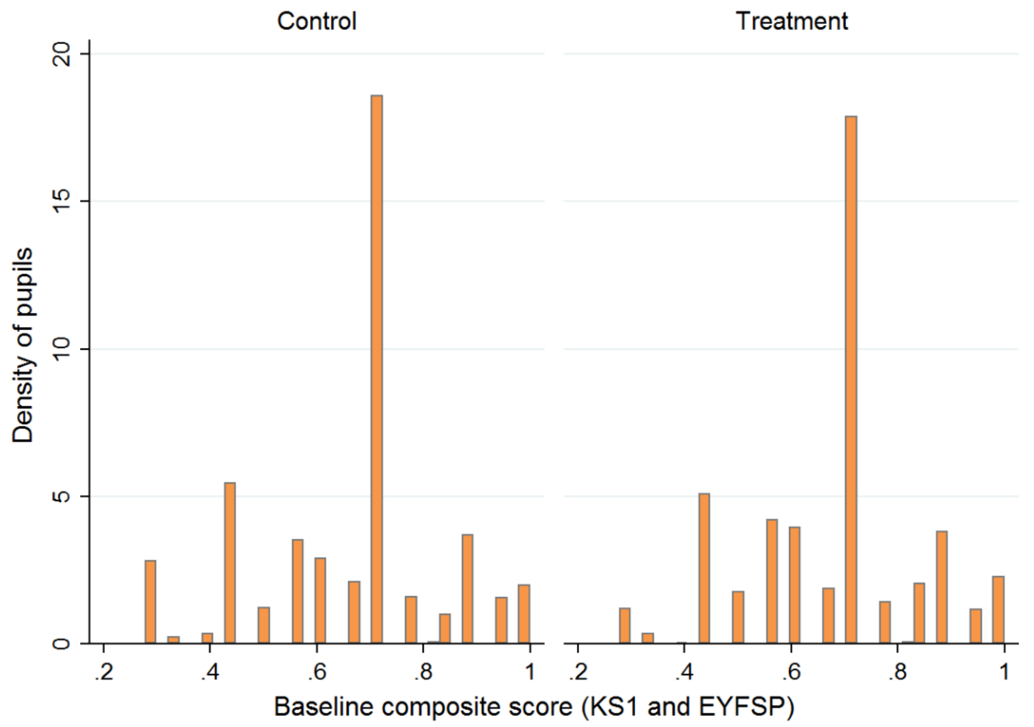
Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	**MDES**	**Attrition**					
5	Randomised design	<= 0.2	0–10%					
4	Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)	0.21–0.29	11–20%	**4**				
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
Threat 1: Confounding	Moderate	The allocation sequence is described clearly. Some moderate imbalance was observed in pre-test at baseline (0.09 on KS1), as well as larger imbalances on school characteristics. Sensitivity analysis controlling for imbalances does not change the headline finding.
Threat 2: Concurrent interventions	Moderate	A range of teaching techniques including White Rose Maths, the Singapore method, Maths Mastery and Explore were reported in control schools, as well as a number of additional mathematics programmes. It is difficult to assess from the IPE the extent to which uptake of these was differential.
Threat 3: Experimental effects	Low	Some control schools described implementing aspects of SDI as part of their usual practice, but this does not seem to stem from active contamination.
Threat 4: Implementation fidelity	Moderate	Implementation data for the key elements of SDI are clearly reported and are aligned with the logic model. Intervention fidelity was variable, with moderate attendance at training and approximately one-quarter of intervention teachers reporting applying a given SDI element occasionally or not at all.
Threat 5: Missing data	Low/Moderate	The amount of missing data was moderate (14% attrition), but approximately balanced across trial arms. Analysis findings were not sensitive to missing data after multiple imputation.
Threat 6: Measurement of outcomes	Low	Test administrators were blinded to treatment allocation and no floor/ceiling effects were observed.
Threat 7: Selective reporting	Low	The trial is registered and pre-specified. Only minimal deviations from SAP occurred (e.g. baseline measure), with reasons clearly justified.

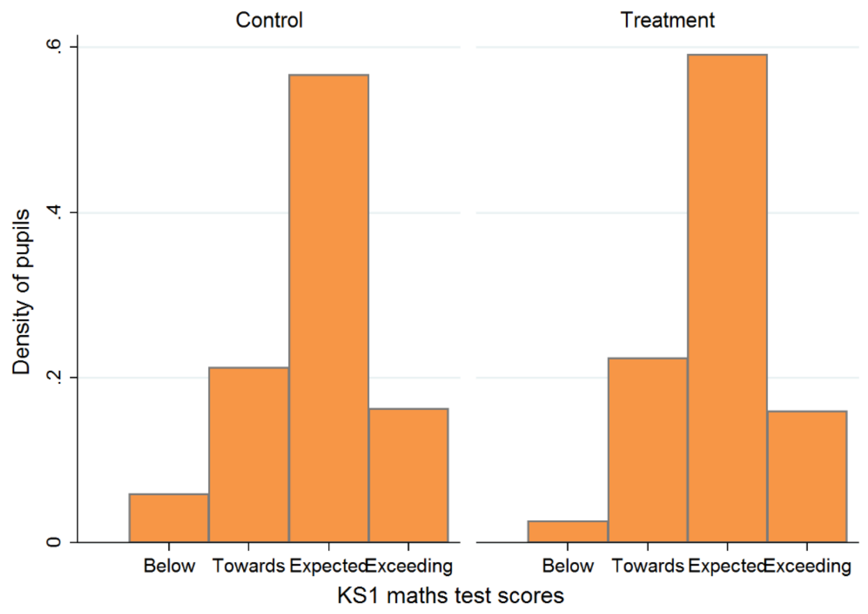
- **Initial padlock score:** [4] Padlocks – Randomised design with MDES at randomisation of 0.28 and attrition of 14.2%
- **Reason for adjustment for threats to validity**: [-1] Padlocks – Three moderate threats to validity
- **Final padlock score:** [3] Padlocks

Appendix C Impact evaluation figures

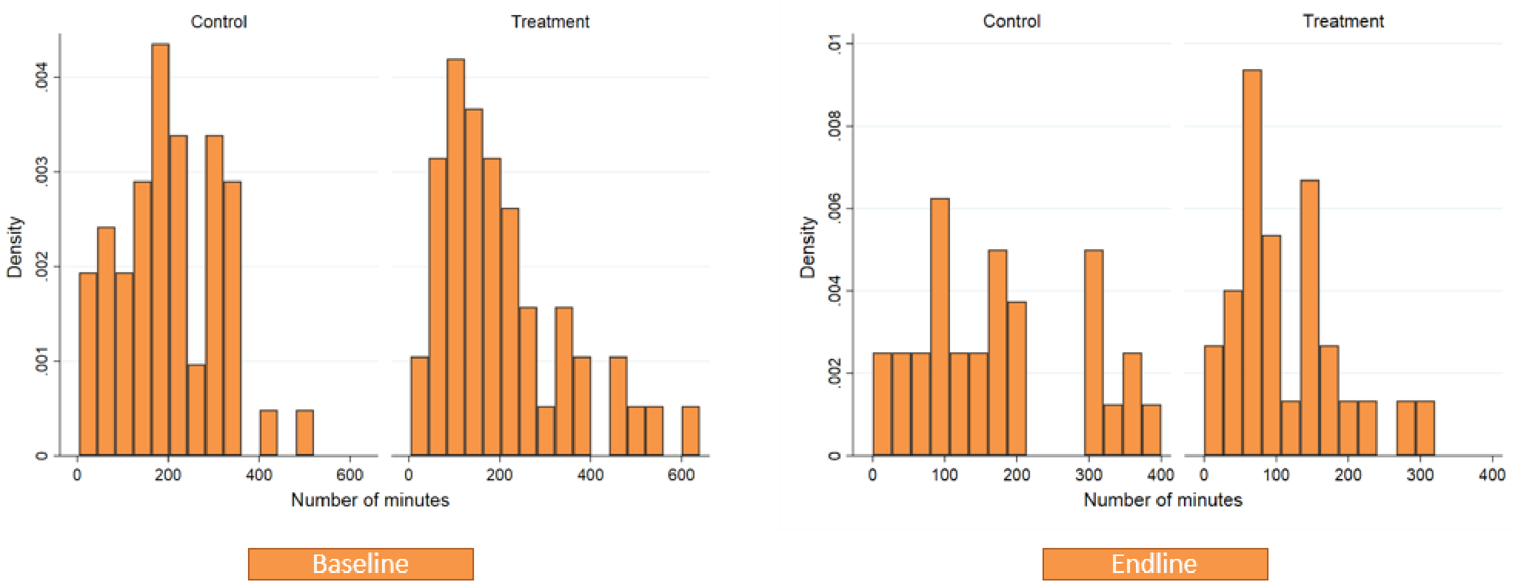
Appendix figure 1: Distribution of the baseline composite score (KS1 and EYFSP)



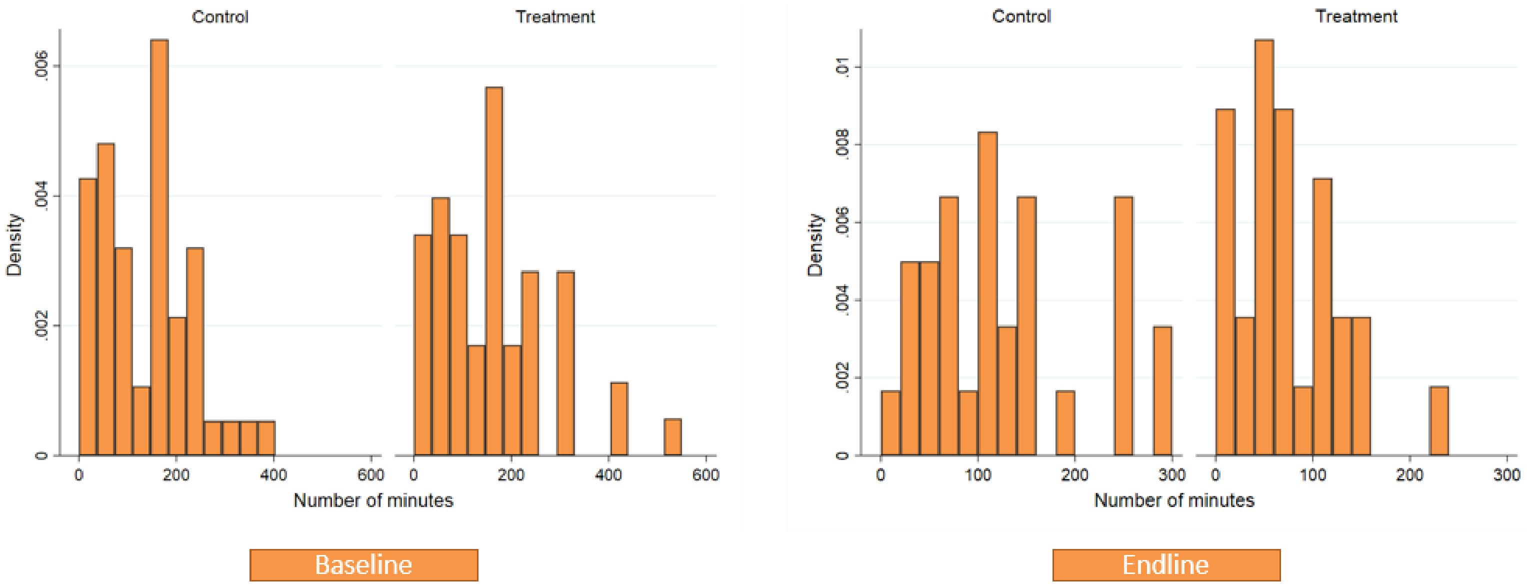
Appendix figure 2: Distribution of KS1 mathematics scores

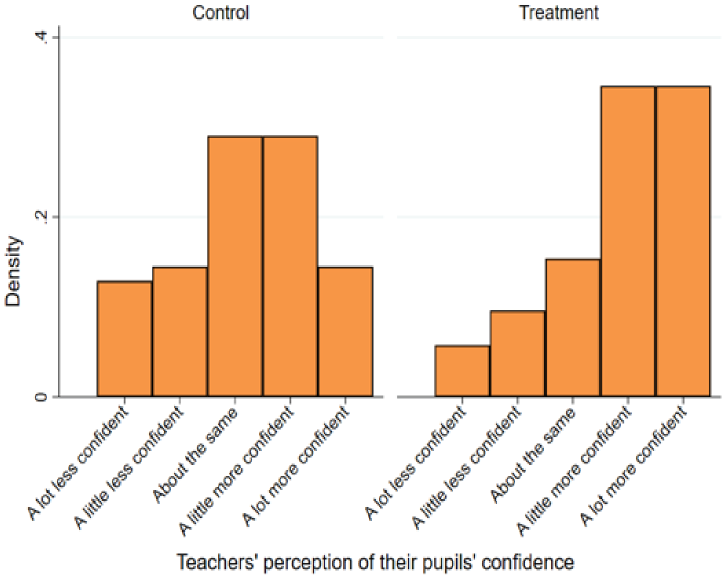


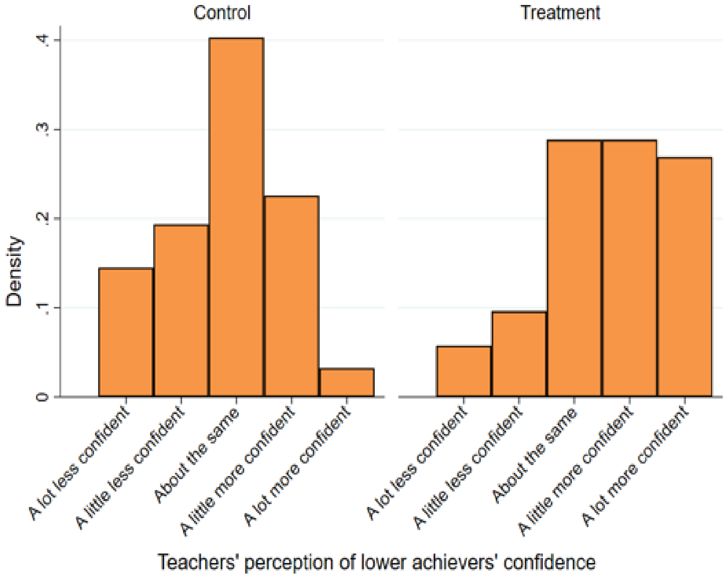
Appendix figure 3: Teacher workload in and outside of class at baseline and endline



Appendix figure 4: Teacher workload outside of class at baseline and endline




Control
Treatment
Density
0
.2
.4
A lot less confident
A little less confident
About the same
A little more confident
A lot more confident
Teachers' perception of their pupils' confidence


Control
Treatment
Density
0
.1
.2
.3
.4
A lot less confident
A little less confident
About the same
A little more confident
A lot more confident
Teachers' perception of lower achievers' confidence

Appendix D Effect size estimation

Appendix table 2: Effect size estimation

			Intervention group		Control group			
Outcome	Unadjusted differences in means	Adjusted differences in means	n (missing)	Variance of outcome	n (missing)	Variance of outcome	Pooled variance	Population variance (if applicable)
Mathematics attainment: *PTM10* raw score (primary analysis)	-0.22	0.15	1,339 (243)	1,172.79	1,491 (225)	1,362.05	1,272.50	N/A
Mathematics attainment: *PTM10* raw score (ever FSM only)	0.55	-0.51	405 (1,177)	452.35	431 (1,285)	403.60	427.21	N/A
Mathematics attainment: *PTM10* raw score (not ever FSM only)	-0.40	0.37	920 (662)	914.94	1,058 (658)	981.49	950.54	N/A
Teachers' workload in class	-62.17	-57.74	27 (29)	5,116.95	29 (33)	11,028.08	8,181.98	N/A
Teachers' workload outside of class	-38.83	-38.87	27 (29)	5,129.13	329 (33)	6,890.58	6,042.47	N/A
Teachers' perception of overall pupils' confidence	-0.69	-0.70	51 (5)	1.36	62 (0)	1.53	2.12	N/A
Teachers' perception of lower achievers' confidence	-0.84	-0.85	51 (5)	1.31	62 (0)	1.11	1.45	N/A
Mathematics attainment: *PTM10* raw score (unadjusted analysis)	-0.06	-0.06	1,357 (225)	1,181.07	1,529 (187)	1,426.80	1,311.26	N/A
Mathematics attainment: *PTM10* raw score (adjusted analysis)	-0.24	0.81	1,325 (257)	1,207.34	1,489 (227)	1,360.05	1,288.15	N/A
Mathematics attainment: *PTM10* raw score (OLS	-0.22	-0.70	1,339 (243)	989.64	1,491 (225)	910.58	947.99	N/A

cluster-robust SE analysis)								
Mathematics attainment: *PTM10* raw score (EYFSP baseline)	-0.26	0.10	1,311 (271)	1,138.52	1,451 (265)	1,328.89	1,238.53	N/A
Mathematics attainment: *PTM10* raw score (imbalanced covariates)	-0.31	0.92	1,275 (307)	1,1205.99	1,451 (265)	1,328.89	1,271.41	N/A
Mathematics attainment: *PTM10* raw score (multiple imputation)	0.35	0.09	1,522 (60)	1184.11	1,716 (0)	1,329.53	1261.18	N/A

Appendix E Intervention description

Further appendices

Further appendix 1: Privacy notice

In line with the new EU General Data Protection Regulation (GDPR), there are certain things that we need to let you, a research participant, know about how your information will be processed. In this privacy notice, we explain the legal basis for data processing, who will have access to your personal data, how your data will be used, stored and deleted, and who you can contact with a query or a complaint.

The National Centre for Social Research (NatCen) is the data controller and data processor for this project. This means that we are responsible for deciding the purpose and legal basis for processing data. For this project, the legal basis for processing data is 'legitimate interest'.
You can contact NatCen with any questions about the research at sameday@natcen.ac.uk.

Who will have access to my personal data?

NatCen is carrying out this research and will have access to pupil, school and teacher information, recordings of interviews, transcripts, charted data, sample files (including contact details), survey responses and anonymised test results. All responses are anonymised before being analysed and archived.
The Yorkshire and the Humber MathsHub, which is part of the Outwood Institute of Education, who developed the Same Day Intervention, will also have access to school and teacher information.
McGowan Transcriptions is the transcription service we use to transcribe our interview and focus group data. They will have access to recordings and transcriptions from all interviews and focus groups. McGowan Transcriptions is on our approved supplier list and compliant with all of our information security policies.
Formara Print+ is the printing company we use to print our materials. They will have access to pupil names and also the names of schools and teachers participating in the project for printing materials. Formara Print+ is on our approved supplier list and compliant with all of our information security policies.
GL Assessment is the provider of the maths tests we are using on this project. It will have access to the names of participating schools and also pupil names, date of birth and Unique Pupil Number. GL Assessment is on our approved supplier list and compliant with all of our information security policies.

How will my data be treated?

The results collected are used for research purposes only. Data gathered from interviews and surveys with teachers and also focus groups with pupils will be used to inform the process study element of the project, while pupils' results from science tests will be used for the impact study element.
At the end of the research, all pupil data will be anonymised before being archived. All personal information, and any other data held on the project, will be securely deleted once the project is complete in April 2021.

Who can I contact with a query or a complaint?

If you have any questions about how your data will be used, please contact the NatCen's Data Protection Officer, Justin Mason.
Under GDPR, you have the right to lodge a complaint with the Information Commissioner's Office. Please go to www.ico.org.uk for more information.

Further appendix 2: Teacher and SLT post-intervention surveys

Teacher survey

The reader should note that the teacher baseline survey consisted of survey part one only.

Overview

This is a short survey for **all current Year 5 teachers in your school**.

Completing this short survey will help us collect the information we need for the evaluation of the Same Day Intervention. The survey should take no more than ten minutes to complete. We know that teachers are incredibly busy, and really do appreciate you taking the time to do this.

Please complete the survey by July 5th 2019

If you have any questions when completing this survey, please contact the NatCen team directly on 0808 169 5668 or email **sameday@natcen.ac.uk**

Survey part one

Please complete all questions.

Q1. Teacher workload {Ask all teachers}

Q1a. Teacher workload:a
This question is about how much time you spend overall on marking Year 5 work from Maths lessons, including time spent marking work inside of lesson time.

During this academic year (2018-19), in an average week, how much time (in minutes) did you spend marking Year 5 work from Maths lessons?

Q1b. Teacher workload:b

This question is about how much time you spend overall on marking Year 5 work from Maths lessons, excluding time spent marking work inside of lesson time.

During this academic year (2018-19), in an average week, how much time (in minutes) did you spend marking Year 5 work from Maths lessons, outside of lesson time?

Survey part two

Q2. Student confidence {Ask all teachers}

The next two questions are about your perception of students' confidence in their Maths abilities. Please compare this year's Year 5 cohort (2018-2019) to last year's Year 5 cohort (2017-2018).

Q2a. As far as you are aware, how does this year's Year 5 cohort that you teach compare to last year's Year 5 cohort?

In regard to their confidence in Maths, compared to last year's cohort, this year's cohort are:

A lot more confident, A little more confident, About the same, A little less confident, A lot less confident

Q2b. Please think of the lowest achieving students in the Year 5 cohort you currently teach. How do they compare to the lowest achieving students from last year's Year 5 cohort?

In regard to their confidence in Maths, compared to last year's cohort, this year's cohort are:

A lot more confident, A little more confident, About the same, A little less confident, A lot less confident

Survey part three

Please complete all questions.

Q3. Same Day Intervention elements {Ask all teachers}

The Same Day Intervention is made up of five key elements. The following questions ask about whether these elements have been a part of Year 5 Maths classes this year.

Q3a. Same Day Intervention elements:a

Thinking about Year 5 Maths classes over the course of the academic year 2018-19, how often have you used Same Day Intervention pedagogical techniques to model new concepts at the start of each Same Day lesson?

Always, Regularly, Occasionally, Not at all

Q3b. Same Day Intervention elements:b

Thinking about Year 5 Maths classes over the course of the academic year 2018-19, how often have you re-structured the Maths lesson to an hour and 15 minutes, including a 15-minute 'pit-stop' to allow the time for marking the Same Day Intervention assignment?

Always, Regularly, Occasionally, Not at all

Q3c. Same Day Intervention elements:c

Thinking about Year 5 Maths classes over the course of the academic year 2018-19, how often have you assigned pupils with an assignment and then marked it whilst pupils were out of the classroom (or otherwise supervised by another member of staff)?

Always, Regularly, Occasionally, Not at all

Q3d. Same Day Intervention elements:d

Thinking about Year 5 Maths classes over the course of the academic year 2018-19, how often was a Teaching Assistant available to support Same Day Intervention classes?

Always, Regularly, Occasionally, Not at all

Q3e. Same Day Intervention elements:e

Thinking about Year 5 Maths classes over the course of the academic year 2018-19, how often did you split the class into two groups based on the results of the assessment, with you teaching the group in need of more support whilst the Teaching Assistant taught the other group?

Always, Regularly, Occasionally, Not at all

The next questions are about the implementation of the Same Day Intervention.

Q3f. Same Day Intervention delivery:f

When did you start delivering the Same Day Intervention in Year 5 Maths classes?

1. September 2018
2. October 2018
3. November 2018
4. December 2018
5. January 2019

Q3g. Same Day Intervention delivery:g

When do you expect to finish delivering the Same Day Intervention? If you stopped delivering Same Day Intervention during the school year, when did this happen?

1. October 2018
2. November 2018
3. December 2018
4. January 2019
5. February 2019
6. March 2019
7. April 2019
8. May 2019
9. June 2019
10. July 2019

Q3h. Same Day Intervention delivering:h

If you stopped delivering Same Day Intervention during the school year, what was the reason for this?

OPEN TEXT <300 CHARACTERS.

Survey part four

Q4. Same Day Intervention Implementation {Ask all teachers}

Q4a. Same Day Intervention implementation:a

The next questions are about your reflections on and opinions of the Same Day Intervention – i.e. the changes made to Year 5 Maths classes.

First, would you recommend Same Day Intervention to another school?

1. Yes

2. No
3. Don’t know

Q4b. Same Day Intervention implementation:b

Please provide an explanation for your answer to Q4a, i.e. if you would recommend Same Day Intervention to another school, why? If not, why not?

:
OPEN TEXT <300 CHARACTERS.

Q4c. Same Day Intervention implementation:c

Do you think your school will continue to deliver Same Day Intervention after this school year?

1. Yes
2. No

Q4d. Same Day Intervention implementation:d

What changes, if any, would you make to Same Day Intervention if you were re-designing it to work best for your school?

OPEN TEXT <300 CHARACTERS.

Survey part five

Q5. Same Day Intervention Staff Time {Ask all teachers}

The next questions are about the time that you have spent on Same Day Intervention.

Q5a. TimeSpent:a (VARLAB: Total time spent preparing for SDI across all teachers)

How many hours in total did you spend on delivering Same Day Intervention during the normal school week? Please express your answer in hours per week, based on an average school week.

This should include all time spent on the Same Day Intervention during the normal school day, including the following: Joint planning – e.g. Same Day Intervention delivery and timetabling in department meetings; Individual planning – e.g. re-designing lesson plans to accommodate Same Day Intervention or preparing diagnostic tasks; Promoting Same Day Intervention throughout the school – e.g. in assemblies or all-staff meetings. Please do not include time in training or time spent outside school hours.

NUMERIC <0..1000> hours

Q5b. TimeSpent:b (VARLAB: Total time spent preparing for SDI across all teachers)

How many hours in total did any TA assisting your class spend on delivering Same Day Intervention during the normal school week? Please express your answer in hours per week, based on an average school week. If you are unsure, please provide your best estimate.

This should include all time spent on the Same Day Intervention during the normal school day, including the following: Joint planning – e.g. Same Day Intervention delivery and timetabling in department meetings; Individual planning – e.g. re-designing lesson plans to accommodate Same Day Intervention or preparing diagnostic tasks; Promoting Same Day Intervention throughout the school – e.g. in assemblies or all-staff meetings. Please do not include time in training or time spent outside school hours.

NUMERIC <0..1000> hours

Q5c. TimeSpent:b (VARLAB: Total time spent preparing for SDI across all teachers)

How many hours in total did you spend on delivering Same Day Intervention outside the normal school day? Please express your answer in hours per week, based on an average school week.

This should include all time spent on the Same Day Intervention outside of normal school hours, including the following: Joint planning – e.g. Same Day Intervention delivery and timetabling in department meetings; Individual planning – e.g. re-designing lesson plans to accommodate Same Day Intervention or preparing diagnostic tasks; Promoting Same Day Intervention throughout the school – e.g. in assemblies or all-staff meetings. Please do not include time in training or time spent inside school hours.

Q5d. TimeSpent:b (VARLAB: Total time spent preparing for SDI across all TAs)

How many hours in total did any TA assisting your class spend on delivering Same Day Intervention outside the normal school day? Please express your answer in hours per week, based on an average school week. If you are unsure, please provide your best estimate.

This should include all time spent on the Same Day Intervention outside of normal school hours, including the following: Joint planning – e.g. Same Day Intervention delivery and timetabling in department meetings; Individual planning – e.g. re-designing lesson plans to accommodate Same Day Intervention or preparing diagnostic tasks; Promoting Same Day Intervention throughout the school – e.g. in assemblies or all-staff meetings. Please do not include time in training or time spent inside school hours.

End of survey

SLT Survey

This is a short survey for the **Head Teacher or senior staff member involved in Same Day Intervention**.

Completing this short survey will help us collect the information we need for the evaluation of the Same Day Intervention. The survey will take around 5 minutes to complete; however, you might need to check your records to provide us with accurate information about the extra time and costs required for this intervention. These surveys are vital elements of the evaluation of the Same Day Intervention. No individuals or schools will be identified in the published evaluation report. The aim of these surveys is to understand how successful the intervention has been and how it can be improved in the future. Please answer as accurately as you are able to. We know that school staff are incredibly busy, and really do appreciate your time.

Please complete the survey by July 5th 2019

If you have any questions when completing this survey, please contact the NatCen team directly on 0808 169 5668 or email **sameday@natcen.ac.uk**

Survey part one

{Ask all}

SDICPD (VARLAB: CPD activities) (Ask all)

The next question is about the continuing professional development (CPD) that Year 5 maths teachers received over the 2018/19 academic year.

Q1a. Please complete the table below with the number of days of CPD Year 5 maths teachers received during this school year (2018/19). If there were no CPD days 2018/19, please enter ‘0’.

Please exclude any potential CPD training days teachers received for the Same Day Intervention in 2018-2019.

	2018/19
Mathematics CPD (total days) Other CPD (total days)	

Survey part two

COSTS

The next questions are about the financial costs incurred by your school in implementing Same Day Intervention. These should not include lesson preparation and teaching time.

Q3a. Training costs

What was the cost of attending Same Day Training? This should include travel and subsistence for all staff that attended training, plus any administration costs incurred.

This should include costs for teachers, teaching assistants and head teachers who attended the training or the optional ‘surgery workshops’. If you are unsure, please provide your best estimate.

Q3b. Teaching Assistant costs

To be implemented as intended, each Same Day Intervention class requires the support of a Teaching Assistant (TA). Did the school need to hire additional TAs in order to staff Same Day Intervention? If so, please estimate the cost to your school.

This should not include costs for TAs that were already employed by your school or would have been needed regardless of Same Day Intervention. If you are unsure of the exact cost, please provide your best estimate.

Q3c. Other costs

What were the costs of printing materials or photocopying for Same Day Intervention classes? If you are unsure, please provide your best estimate for total cost for the entire duration of the intervention.

TEACHER COVER

Q4a. Cover arrangements

How many days of cover were needed to allow school staff to attend Same Day Intervention training?

This should include time for cover for teachers, TAs and head teachers to attend the training or the optional 'surgery workshops'. If you are unsure of the exact number of days, please provide your best estimate.

MATHSPROG (VARLAB: maths programmes) (Ask all)

The following question is about any maths intervention (other than Same Day Intervention) for Year 5 students that you might have implemented during 2018/2019

Q1b. Other than Same Day Intervention, has your school delivered any maths interventions for Year 5 students during 2018/2019? Please select all that apply

- Catch Up Numeracy
- Digital Feedback in Primary Maths
- Numbers Count
- 1st Class@Number
- Success@Arithmetic
- Numeracy Ninjas
- Pre-teaching Intervention
    - On Track Maths
    - Mission Maths
    - Dynamo Maths
    - Third Space Learning
- Other (Open answer)
- None of the above

End of survey

Further appendix 3: Memorandum of Understanding

**Agreement to participate in the evaluation of
Same Day Intervention**

Thank you for participating in this evaluation of Same Day Intervention. The purpose of the following agreement is to outline the commitments of all parties involved.

Please sign both copies, retaining one and returning the second copy to [NAME OF CONTACT] at [PROJECT DELIVERY ADDRESS/EMAIL] by [ADD DATE]

School Name: ____

School Postcode: _____________ Head teacher Name: ____________________________

In order for your school to be able to deliver Same Day Intervention each Year 5 class must have a Teaching Assistant available to support <u>all</u> maths lessons during the 2018-2019 academic year.

To participate in Same Day Intervention your school will need to have completed the following by 27th April 2018:

- **Sent parental opt-out letters to all parent/carers of Year 4 pupils and left a two-week period for any responses**
- **A pupil upload form, which requires the school to provide the details of all current Year 4 pupils**
- **A short teacher survey – to be completed by all teachers due to teach Year 5 classes in the 2018-2019 academic year**
- **A short school survey – to be completed by the head teacher or project lead**

1. Aims of the Evaluation

The aim of this project is to evaluate the impact of Same Day Intervention.

Same Day Intervention has been developed by the Yorkshire and the Humber Maths Hub. The intervention introduces a new approach to Maths teaching for Year 5 pupils, with the aim that all pupils succeed and become resilient and confident mathematicians. In Same Day Intervention, the structure of the lesson includes a 'progress pit-stop' placed in the middle, during which teachers can assess and group children for further teaching.

By participating in this research, you will make an important contribution in helping to understand methods of teaching that prevent the attainment gap.

2. Evaluation Design

The research design is a randomised controlled trial and it will involve 120 primary schools in Yorkshire and the Humber and surrounding areas. Half the schools will be randomly assigned, by computer, to the treatment group and will deliver Same Day Intervention in 2018-19. The other half will be allocated to the control group and continue their usual teaching. All schools (control and treatment) are required to involve all Year 5 pupils in a Maths test at the end of the 2018-19 academic year, alongside other evaluation activities which are detailed in Section 3 of this document.

The control schools can then offer Same Day Intervention to the next cohort of Year 5 pupils if they would like to. The evaluation is being conducted by an independent research team from NatCen Social Research.

Random assignment of schools to control and treatment is essential to the evaluation as it is the best way of identifying what effect the project has on children's outcomes. It allows the research team to compare progress made by children in treatment and control schools. It is important that schools understand and consent to the random allocation process.

3. School Responsibilities

3.1. Prior to the trial:

- <u>Pupil recruitment</u>: Recruitment of all Year 5 classes within the school. **Schools where any Year 5s are taught in mixed year group classes cannot be included in the trial.**
- <u>Have dedicated Same Day Intervention teacher(s)</u>

- Have a named Same Day Intervention Co-ordinator: The co-ordinator will play a key role in delivering Same Day Intervention and supporting the NatCen research team to conduct the evaluation tasks. They will be the main point of contact when NatCen is scheduling evaluation tasks. A member of the senior management team at each school also needs to be appointed and named as a key contact.
- Consent to randomised allocation and commit to the outcome: whether assigned to the treatment or control group.

3.2. During the trial:

- Control schools are requested not to run Same Day Intervention with Year 5 pupils during the 2018-19 academic year: control schools can deliver their planned activity with Year 5 the following academic year (2019-2020) but are asked not to deliver Same Day Intervention with the 2018-19 Year 5 cohort.
- Teaching Assistant (TA) supporting during all maths lessons in Year 5 classrooms: The Teaching Assistant plays an essential role in the intervention. They will work with pupils who have grasped the concept in focus to undertake further tasks.
- Attend training sessions: All Year 5 teachers due to deliver Same Day Intervention in the academic year 2018-19 will be required to attend three whole-day training sessions and at least two Open Classroom events to observe and learn how to implement Same Day Intervention. The Headteacher will be required to attend one Open Classroom event to see Same Day Intervention in action.
- Adapt Maths lessons to include break in lesson: All Maths lessons delivered to Year 5 pupils in the intervention schools will require 15 minutes of non-teaching time. This can be facilitated by a break in the lesson via an assembly or other activity, or the TA can teach the class for 15 minutes. This 15-minute 'pit-stop' is essential so that teachers can mark pupils' work and group pupils into those who understood concepts being taught and those who did not.
- Keep a record of additional costs incurred while running Same Day Intervention, if allocated to the treatment group.

3.3. Evaluation tasks:

- Commit to completion of all pre- and post-trial tasks and liaise with research staff to find times and dates for these tasks: evaluation tasks will be conducted with school staff and pupils. The tasks required for the control and treatment groups are summarised in the table and described in more detail below.

Evaluation Activity	Control Group	Treatment Group
1. Distributing parental data processing objection letter letters to parents of Year 4 pupils **Parents of all Year 4 pupils (who will be moving into Year 5 in the academic year 2018-19) must receive a letter enabling them to object to processing their child's data for the of the project evaluation.**	✓	✓
2. Complete teacher survey and submit pupil information **Same Day Intervention teachers in all schools will be asked to complete a short pre-trial questionnaire between February and April 2018. Schools will also be asked to provide pupil level information in between February and April 2018.**	✓	✓
4. Case studies in nine schools **Case studies will be undertaken in nine treatment schools. These schools will be invited to take part in a visit from the NatCen research team during the Spring term of 2018-19.** **The visit will involve a NatCen researcher conducting face-to-face interviews with the staff involved in Same Day Intervention, and discussion groups with pupils in Year 5 who have participated in the intervention.**		✓
5. Telephone interviews in five schools **A small number of staff at control schools will be asked to take part in a telephone interview during the Spring term of 2018-19.**	✓	
6. School end of trial survey **All participating teachers will be asked to complete a short post-trial survey in June 2019.**	✓	✓
7. Pupil end of trial survey and Maths assessment **All schools will take part in a telephone call with NatCen staff to schedule and plan a Maths assessment for pupils. NatCen staff will visit the school and facilitate the completion of the one-hour online assessment by all eligible pupils**	✓	✓

in June 2019. As part of this, the school will ensure that a teacher is present on the day of the assessment to support in managing the behaviour of pupils.		

4. Yorkshire and the Humber MathsHub

- Provide information about Same Day Intervention, and training on how to deliver Same Day Intervention in the classroom. For schools allocated to the treatment group, this will begin in June 2018 and for schools allocated to the control group who can choose to offer Same Day Intervention to the next cohort of Year 5 pupils, this will begin in September 2019.
- Be the first point of contact for any questions about the trial.
- Provide on-going support to the school throughout the trial. You can contact the team via email: oie@outwood.com

5. NatCen responsibilities

- Provide materials to enable schools to undertake the parental opt-out process.
- Conduct the random allocation of schools to treatment or control groups.
- Provide templates to schools to help them record any additional costs incurred by participating in the trial.
- Be a secondary point of contact for questions specifically about the evaluation.
- Facilitate the Maths assessment that is due to take place with all Year 5 pupils in June 2019.
- Collect and analyse the data from the project to estimate the impact of the treatment, including pupil and school data, assessment results, interviews with school staff and discussion groups with pupils.
- Write a report published by EEF, on the findings of the project and disseminate research findings (through publication on NatCen website) in which no school or pupil will be identified.

6. Payment

- Control schools will receive £1,000: This will be a staged payment with £200 at the start and £800 later on when all evaluation tasks are complete.

7. Use of data

All data, including pupils' test responses and any other pupil data, will be treated with the strictest confidence. Pupil assessments will be administered by NatCen Social Research. Named data will be matched with the National Pupil Database and shared with NatCen, Yorkshire and the Humber Maths Hub, the Department for Education, EEF, EEF's data contractor FFT Education and in an anonymised form to the UK Data Archive. No individual school or pupil will be identified in any report arising from the research.

We commit to the evaluation of Same Day Intervention as detailed above

Head teacher/Senior leader

Name:

Signature: ______________________________ Date: ___________

Same Day Intervention Year 5 teacher

Name

Signature: _______________________________ Dates_____________

Further appendix 4: Parent information letter

Dear Parent/Carer,

RE: Evaluation of Same Day Intervention

I am writing to tell you about an exciting new maths programme your child's school may be participating in next year – **Same Day Intervention (SDI).**
Same Day Intervention introduces a new structure and approach to Maths teaching for Year 5 pupils. The new structure involves the teacher demonstrating a topic, before children complete five or six questions related to that topic. There is then a 15-minute 'pit-stop' in the lesson, during which teachers mark pupils' work. The second part of the lesson involves the teacher providing intensive support for pupils who need further help to understand the topic, while pupils who were able to grasp the concept are given tasks to further their learning. This is different to the standard Maths lesson where pupils' work is marked at the end of the lesson by the teacher. The aim of this new structure is to ensure that all pupils can attain a good level of mathematics understanding and that no pupil gets left behind. This approach is based on successful methods from Shanghai, which could improve learning for all children.

NatCen Social Research is evaluating Same Day Intervention to understand if it makes a difference to pupils. NatCen will be randomly assigning schools to either an 'intervention' or a 'control' group. Schools in the 'intervention' group will run Same Day Intervention in 2018-19, while schools in the 'control' group will continue with the current best practice in maths teaching. This means there is a 50/50 chance that your child's school will be offering Same Day Intervention.

We are asking for your permission to obtain information about your child from the school, including their name, date of birth and Unique Pupil Number. Information provided will be shared with a small group of researchers at NatCen Social Research. We will use the information provided to link with the National Pupil Database (which is held by the Department for Education). Anonymised data will be shared with the Department for Education, the Education Endowment Foundation who are funding SDI and EEF's data contractor FFT Education. Once your child's information is included in the data set, the data will be anonymised and no one will be able to identify individual pupils.

All pupils involved in the evaluation, regardless of whether they are in the intervention or control group, will undergo an assessment in June 2019. This will take one hour to complete. The assessments will be invigilated and collected by NatCen Social Research, processed by GL Assessment, and the results will only be accessible to the researchers at NatCen working on the project. The data will be kept confidential, in accordance with the 1998 Data Protection Act and the 2018 General Data Protection Regulation. Only group results of the programme evaluation will be published. Your child's name and the name of the school will not be used in any report arising from the research.

Some schools taking part in SDI will be involved in a case study. Researchers from NatCen Social Research will visit case study schools to observe the Maths lesson and also speak with some of the pupils in the class. We will contact you separately to let you know if your child will be involved in this.

Please complete the slip below if you **do not** want your child's school to share your child's information with the evaluators and wish your child to opt out of the evaluation.

If you would like more information about this evaluation, you can contact the research team at NatCen, using the telephone number or email address below.

If you are happy with everything in this letter you do not need to do anything.

Malen Davies
Senior Researcher
NatCen Social Research

Tel: 0808 169 5668
Email: sameday@natcen.ac.uk
Website: http://www.natcen.ac.uk/taking-part/studies-in-field/evaluation-of-same-day-intervention/

Same Day Intervention: parental objection to data processing form

I do **not** want my child to take part in the evaluation of Same Day Intervention, including all the research activities described above:

Your child's name**…………………………………………………………………………….**

Your full name**……………………………………………………………………………..**

Your signature…………………………………………….

Date………………………………....

Please return this slip to your school within **two weeks** of receipt of this letter if you wish to **opt-out** of the evaluation of Same Day Intervention.

Further appendix 5: Randomisation syntax

```
clear
cap log close
set more off
*version 14.1

log using "xxx", replace

*********************************
*P12509
*Same Day - randomisation
*Daniel Phillips 11/07/18
*********************************

*************************
*Anonymisation of dataset
*************************
// open data
cd "xxx"
import excel using "Plus2randomisation.xlsx", sheet ("Random2") firstrow clear case(lower)

//generate unique setting ID for evaluation
set seed 9990/* from random.org, random number between 1 and 10,000 generated on 23.08.18*/
display "seed" %10.0f c(seed)
generate double r = 0
replace r = runiform() /* assigns random number to each setting*/
sort r /* sorts settings by random number from smallest to largest*/
generate settingID = _n /* generates a case ID for each setting in ascending order*/
label variable settingID "Setting ID number assigned for evaluation"

//save dataset w/ all IDs
sort hubii
drop r
save P12509SampleWithSettingIDs, replace
export excel using "P12509SampleWithSettingIDs_2.xlsx", firstrow(var) replace

// drop identifying info
drop laestno urn schoolname
save P12509SampleforRandomisation, replace
export excel using "P12509SampleforRandomisation_2.xlsx", firstrow(var) replace

*************************
*Randomisation
*************************

// randomisation setup
egen hb=group(hubii),label missing
g _assignment = 0
label define _alb 0 "No assignment(error)" 1 "Treatment" 2 "Control"
label values _assignment _alb

*number of schools per hub
ta hb, m

//randomisation
display "seed" %10.0f c(seed)
generate double r = 0
replace r = runiform()

*assignment
sort hb r
by hb: replace _a = 1 if _n<=_N/2
by hb: replace _a = 2 if _n>_N/2

* last replication correction
/* for strata (hubs) with an odd number of schools
randomly allocate last school to either treatment and control condition
to minimise imbalance in treatment allocation across the sample
*/
```

```
generate double lrc = 0
label define _lrclb 0 "No correction required" 1 "LRC"
label values lrc _lrclb

by hb: replace lrc = 1 if (_n==_N) & mod(_n,2)==1
replace _a=0 if lrc==1
generate lrcr = runiform() if lrc==1
replace _a = 1 if lrcr<=0.5
replace _a = 2 if lrcr>0.5 & lrcr<=1

// check treatment allocation balance
ta _a, m

// display the state of the random-number generator
/*allows continuation of randomisation from last place left off
following 1st wave of random number generation
*/
display "seed" %10.0f c(seed)

//save output
sort hb _a
drop hb r lrc lrcr

export excel using "P12509SampleWithAssignment_2.xlsx", firstrow(var) replace
save P12509SampleWithAssignment2, replace

****************************************
*Merge assignment with identifying info
****************************************
use "P12509SampleWithSettingIDs2", clear
merge 1:1 settingID using "P12509SampleWithAssignment2.dta"
ta _merge, m
drop _merge

******************************
*save final randomised dataset
******************************
save "xxx", replace
export excel "xxx", firstrow(var) replace

******************************
*end of do file
******************************
log close _all
```

Further appendix 6: Effect-size calculation syntax

```
*Analysis syntax:
use "EYFSP_KS1_2Census_GL_merged_as_randomised.dta", clear

* Complete case sample
keep if !missing(GL_TotalRawScore, treat, KS1_MATH_OUTCOME_cat, hub, SchoolID, GL_Class)

*****
* 1. 3-Level Hierachical Linear Model
*****

mixed GL_TotalRawScore treat ib3.KS1_MATH_OUTCOME_cat i.hub ||  SchoolID: ||  GL_Class:,mle
variance // Highest level is stratification (Fixed Effects), then schools, then pupils

*****
* 2. Treated and Control group sample sizes at pupil level
*****

ta treat, matcell(x)
matrix list x
local n1=x[2,1]  //Treated
local n2=x[1,1]  // Control
local N=x[1,1]+x[2,1] //Total
display _column(20) `n1'     _column(40) `n2'    _column( 60) `N'

*****
* 3. Adjusted means difference
*****

*mixed GL_TotalRawScore treat baseline_rescale i.hub ||  SchoolID: ||  GL_Class:,mle  variance
matrix define k=r(table)
matlist k
local Yadj_treated=k[1,1]  + k[1,9]
local Yadj_control=k[1,9] // update coordinates
local pvalue_adj=k[4,1]
local adj_diff=`Yadj_treated'-`Yadj_control'
display  _column(20)  `Yadj_treated' _column(40) `Yadj_control' _column(60) `pvalue_adj'

************************
* 4. Unadjusted mean difference
************************

ttest GL_TotalRawScore, by(treat)

local Yunadj_treated=r(mu_2)
local Yunadj_control= r(mu_1)
local unad_diff=`Yunadj_treated'-`Yunadj_control'
local Yunadj_treated_LB=33.65456
local Yunadj_treated_UB=35.25955
local Yunadj_control_LB= 33.91557
local Yunadj_control_UB= 35.4352
local pvalue_unadj=r(p)

display  _column(20)  `Yunadj_treated_LB' _column(40) `Yunadj_treated_UB' _column(60)
`Yunadj_control_LB' _column(80)`Yunadj_control_UB'
display  _column(20)  `Yunadj_treated' _column(40) `Yunadj_control'  _column(60) `unad_diff'
_column(80)`pvalue_unadj'

*****
* 5. Pooled standard deviation
*****
/*
Pooled standard deviation = \_/((n1-1)*sd1^2+(n2-1)*sd2^2)/(n1+n2-2)
*/
* EEF Preference is to use the unconditional variance
* Treated
mixed GL_TotalRawScore || SchoolID: || GL_Class: if treat==1, mle variance
matrix define z1 = r(table)
matlist z1
```

```
local se1 = z1[2,1] // takes the standard error
local sd1 = `se1'*sqrt(`n1') // convert the std.error to std.deviation

* Control
mixed GL_TotalRawScore || SchoolID: ||  GL_Class: if treat==0, mle variance
matrix define z2 = r(table)

local se2 = z2[2,1] // takes the standard error
local sd2 = `se2'*sqrt(`n2') // convert the std.error to std deviation

* Estimate pooled standard deviation:

display _column(20)`sd1' _column(40)`sd2'
local pooled_sd = sqrt(((`n1'-1)*((`sd1')^2))+((`n2'-1)*((`sd2')^2)))/(`n1'+`n2'-2))
display `pooled_sd'

*****
* 6. Number of schools
*****

preserve
keep if treat==1
keep SchoolID
duplicates drop
count
local m_t=r(N)
restore

preserve
keep if treat==0
keep SchoolID
duplicates drop
count
local m_c=r(N)
restore

display _column(20) `m_t' _column(40) `m_c'

*****
* 7. Number of class
*****

preserve
keep if treat==1
keep GL_Class SchoolID
duplicates drop SchoolID GL_Class , force
count
local p_t=r(N)
restore

preserve
keep if treat==0
keep GL_Class SchoolID
duplicates drop SchoolID GL_Class , force
count
local p_c=r(N)
restore

local P=`p_c'+`p_t' //Total
display _column(20) `p_t' _column(40) `p_c' _column(60) `P'

*****
* 8. Estimation of ICC's
*****

mixed GL_TotalRawScore treat ||  SchoolID: ||  GL_Class:,mle  variance
estat icc
local rho_s=23.16167 /(23.16167 + 4.479969   + 194.4998)
```

```
local rho_c= 4.479969   /(23.16167 + 4.479969  + 194.4998)

display _column(20) `rho_s' _column(40) `rho_c'

*****
* 9. Nu
*****
sort SchoolID
by SchoolID: gen number_pupils=_N
gen number_pupils2 = number_pupils^2

preserve
keep if treat==1
keep SchoolID number_pupils2
duplicates drop
egen temp=total(number_pupils2)
sum temp
local sum_nt_squared=r(mean)
restore

preserve
keep if treat==0
keep SchoolID number_pupils2
duplicates drop
egen temp=total(number_pupils2)
sum temp
local sum_nc_squared=r(mean)
restore

local nu_1 = (`n2' * `sum_nt_squared')/(`N'*`n1')
local nu_2 = (`n1' * `sum_nc_squared')/(`N'*`n2')

local nu=`nu_1'+`nu_2'

display _column(20) `nu_1' _column(40) `nu_2'  _column(60) `nu'

***********
* 10. Pu
***********

sort SchoolID GL_Class
by SchoolID GL_Class: gen number_pupils_class=_N
gen number_pupils_class2 = number_pupils_class^2

preserve
keep if treat==1
keep SchoolID GL_Class number_pupils_class2
duplicates drop
egen temp=total(number_pupils_class2)
sum temp
local sum_pu_squared_t = r(mean)
restore

preserve
keep if treat==0
keep SchoolID GL_Class number_pupils_class2
duplicates drop
egen temp=total(number_pupils_class2)
sum temp
local sum_pu_squared_c = r(mean)
restore

local pu_1=(`n2'*(`sum_pu_squared_t'))/(`N'*`n1')
local pu_2=(`n1'*(`sum_pu_squared_c'))/(`N'*`n2')

local pu=`pu_1' +`pu_2'

display _column(20) `n1' _column(40) `n2' _column(60) `N'
```

```
display _column(20) `sum_pu_squared_t' _column(40) `sum_pu_squared_c'

display _column(20) `pu_1' _column(40) `pu_2' _column(60) `pu'

***********
* 11. N tilda
***********

local n_tilda=(`n1'*`n2')/(`n1'+`n2')

display _column(20) `n_tilda'

*************************************
* 12. J bias correction
*************************************
local J=1-(3/(4*(`n1'+`n2'-2)-1))

display _column(20) `J'

*************************************
* 13. Hedges g
*************************************

local cohend = ((`Yadj_treated'-`Yadj_control')/(`pooled_sd'))

display _column(20) `cohend'

local correction= 1-(((2*(`pu'-1)*`rho_s')+(2*(`nu'-1)*`rho_c'))/(`N'-2))
display _column(20) `correction'
local correction_sqrt=sqrt(`correction')

local hedgesg=`J'*`cohend'*`correction_sqrt'

display _column(20) `hedgesg'

* clustering correction
local cluster_correction2=1-(((2*(`pu'-1)*`rho_s')+(2*(`nu'-1)*`rho_s'))/(`N'-2))
local cluster_correction=sqrt(`cluster_correction2')
display _column(20) `cluster_correction'

* Hedges'g effect size
local hedgesg=`J'*((`Yadj_treated'- `Yadj_control')/`pooled_sd')*`cluster_correction'


*************************************
* 14. Estimate of the variance of D
*************************************

local V_g1 = (1+((`pu'-1)*`rho_s') + ((`nu'-1)*`rho_c'))/(`n_tilda')
local V_g2=(`hedgesg'^2)/(2*(`m_t'+`m_c'-2))

local V_g=`V_g1' +`V_g2'
display  _column(20) `V_g1' _column(40) `V_g2' _column(60) `V_g'


*************************************
* 15. Hedges' g & CI
*************************************

local g_lb = `hedgesg'-(1.96*sqrt(`V_g'))
local g_ub = `hedgesg'+(1.96*sqrt(`V_g'))

display  _column(20) `hedgesg' _column(40) `g_lb' _column(60) `g_ub'

log close
```

Further appendix 7: Topic guide, teachers in intervention school

Evaluation of the Same Day Intervention

Topic guide

Interviews with teachers – case study schools

Start recording and ask permission to start recording. If they don't agree to recording, take handwritten notes.

Context

Aim: gather information on school context and respondent background.

Background information

- Overview of school
    - Contextual factors – size, rural/urban
    - Number of Year 5 classes/teachers
    - Number of TAs for Year 5
- Overview of respondent's role and responsibilities
- Respondent's teaching experience
    - Number of years teaching
    - Years teaching Year 5
    - Years in current school
- Overview of staff involved in Same Day Intervention
    - Any staffing changes or staffing gaps since the start of year

Training

Aim: gather perceptions of usefulness of training received for delivering Same Day.

- Attendance at Open Classroom

  If no:
    - Reason why
    - Whether problematic for intervention delivery
    - Was the training cascaded
    - How was training cascaded

  If yes:
    - Usefulness
    - What liked about training

  - Anything did not like
  - Views on resources given
    - Adequacy of resources
  - Suggested improvements to Open Classroom

- Attendance at other training days (day 2 and day 3)

  If no:
  - Reason why
  - Whether problematic for intervention delivery
  - Was the training cascaded
  - How was training cascaded

  If yes:
  - Usefulness
  - What liked about training
  - Anything did not like
  - Views on resources given
    - Adequacy of resources
  - Suggested improvements to training

- Whether training cascaded to TAs

  If no:
  - Reason why
  - Any implications on delivery

  If yes:
  - What training has been cascaded
  - How was training cascaded
  - Whether necessary
  - Usefulness

- Attendance at surgery session

  If no:
  - Reason why
  - Whether problematic for intervention delivery

  If yes:
  - Usefulness
  - Whether session helped resolve queries
  - Any further clarification needed

  - Any other feedback
  - Any changes to intervention delivery after session
  - Suggested improvements to session

If any changes to staff since the start of the year

  - Training cascaded to new members of staff involved in Same Day Intervention
  - If yes – easy / difficult? Why

- Overall views on training and resources
  - What resources were given
  - More or less training needed
  - Adequate resources
- Whether felt ready to deliver Same Day Intervention
  - If no, what other support was required

Delivery

Aim: explore how Same Day Intervention was set up and is being delivered, key challenges and successes factors and further support or resources needed.

Experience of set-up and planning

- Any costs involved in implementing Same Day Intervention
  - Resources for teachers
    - One-off or ongoing
  - Travel costs to attend training
  - Costs to cover training (e.g. supply teachers)

- Experience of:
  - Scheduling of Same Day Intervention into timetable
  - Adding the 15-minute pit-stop into timetable
    - How easy / difficult

- Experience of lesson planning for Same Day Intervention
  - How easy / difficult
  - Any changes over time
  - Whether received support from others (teachers in or outside school)

Experience of delivery

- Overall experience of delivering Same Day Intervention so far
  - Whether delivering Same Day Intervention in every maths lesson

        - If no, reasons for this
    - Aspects working well
        - TA support
    - Aspect working less well / key challenges
        - Whether resolved challenges
    - Any changes to delivery since September
        - If yes, what have these been
        - Why were implemented
- Views on each element of Same Day Intervention (if not already covered)
    - Modelling (i.e. exercise used to demonstrate topic being taught)
    - Diagnostic tasks (i.e. task used to diagnose which group pupil falls into)
    - Pit-stop (i.e. 15-minute break where teachers must mark pupils' task)
    - Bronze, Silver and Gold task (i.e. tasks pupils do if they have demonstrated they grasp the topic being taught)
    - Same Day Intervention (i.e. group of pupils who receive intensive support to help understand topic being taught)
- Timings of lessons (e.g. first part / second part of lesson)
- Experience of the TA support
        - Usefulness
        - Any issues
- Whether each aspect of Same Day Intervention is being used
        - If not, what is not being used and why
- Support received
    - View on support received from SLT
- Any further support needed
- Any further resources needed to aid delivery
- View on support from other Same Day Intervention schools / teachers

Perceived impacts

Aim: to explore what impacts delivering Same Day Intervention has had on teachers and pupils.

- Perceived impacts for teachers
    - Workload
        - More or less time planning compared to last year
        - More or less time marking compared to last year
    - Confidence
    - Anything else

- Perceived impacts for pupils
    - Maths skills
    - Enjoyment of Maths
    - Self-esteem / confidence in Maths
    - Anything else
- Whether had impact on specific groups
    - Pupils of average ability
    - Lower ability or SEND pupils
- Longer term impacts
    - Teachers
    - Pupils
- Any unexpected impacts
- Any other impacts

Overall summary

- Key challenges
- Key successes
- If a teacher from another school asked about the Same Day Intervention, what would you tell them?
- Would you recommend the Same Day Intervention to another school?
- Whether planning on delivering Same Day Intervention next year?

TURN OFF RECORDER

Ask if any concerns about what they have told us (if concerns, we can redact sections).

Thank participant and close.

Further appendix 8: Topic guide, teachers in control schools

Evaluation of the Same Day Intervention

Topic guide

Interviews with a Year 5 teacher – control schools

Respondent background and school context

Aim: gather information on school context and respondent background.

Overview of school

- Contextual factors –size, rural/urban,
- Number of Year 5 classes/teacher

Overview of respondent's role and responsibilities

- Current role
- Length of time in role
- Reasons for choosing to be part of Same Day project

Delivery of Maths lessons

Aim: to understand normal practices in delivering Maths lessons for Year 5 pupils.

Whether pupils are set / streamed based on ability

Ask teachers to describe typical Maths d for Year 5 pupils

Any particular type of lessons / technique used during lessons

- Maths mastery
- Other techniques
- Whether delivering any aspects of Same Day Intervention
  - Diagnostic (i.e. task used to diagnose which group pupil falls into)
  - Pit-stop (i.e. 15-minute break where teachers must mark pupils' task)
  - Same Day Intervention (i.e. group of pupils who receive intensive support to help understand topic being taught)

Types of resources used

- Purchased textbooks i.e. Power Maths or Maths – No Problem!
- White Rose
- Class secrets
- Internally developed resources
- Any other resources used

Views on approach to teaching Year 5 Maths

- Aspects that work well

        - Resources
        - Techniques
    - Aspects that work less well
    - Whether meet needs of all pupils
    - Any pupils approach has been unsuccessful with

Teacher workload

    - Feelings around marking / planning Maths lessons
        - Main issues / difficulties with marking / planning Maths lessons
        - Higher or lower workload compared to other subjects

    - Approaches that help reduce marking / planning time of Maths lessons
    - Approaches that might increase marking / planning time of Maths lessons

Summary
Aim: to offer an opportunity for any final thoughts/reflections.

Key challenges teaching in teaching Maths

Key elements successes when teaching Maths

Anything else to add

Ask whether have any questions

Further appendix 9: Exemplar SDI materials

Below are examples of:

1. Diagnostic questions used by teachers following the input;
2. The Silver and Gold tasks given to pupils based on their in-class assessment;
3. Planning used by teachers to map out the medium plan (objectives for each lesson) and small steps (breaking each learning objective down into a series of smaller steps);
4. Collaborative lesson planning from the surgery sessions.

1. Autumn Term, Week 8, Lesson 1 diagnostic

1) Which number is not a multiple of 3?
21 3 15 13 18 12

2) Look at these representations.
What multiple is represented?

List the factors.

3) 20 = ☐ x ☐
20 = ☐ x ☐
20 = ☐ x ☐
The factors of 20 are ______

4) Polly is planting potatoes in her garden. She has 24 potatoes to plant and she will arrange them in a rectangular array.

5) Circle the common factor of 7 and 35?
3 6 7 9 35

6) Which two numbers have common factors of 3 and 4?
16 24 30 36

2. Autumn Term, Week 8, Lesson 1 Silver and Gold tasks

Silver

8 is a multiple of 4 and a factor of 16		
6 is a multiple of 3 and a factor of ☐		
☐ is a multiple of 5 and a factor of ☐		
☐ is a multiple of ☐ and a factor of ☐		

Make five numbers

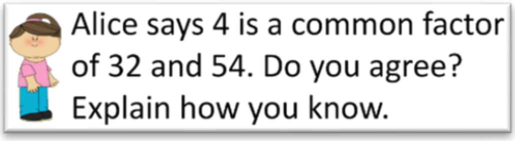
Take ten cards numbered 0 to 9.

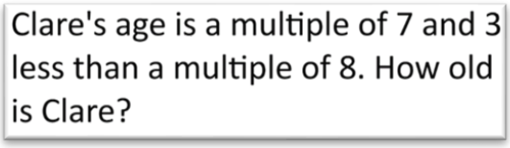
Each time use all ten cards.

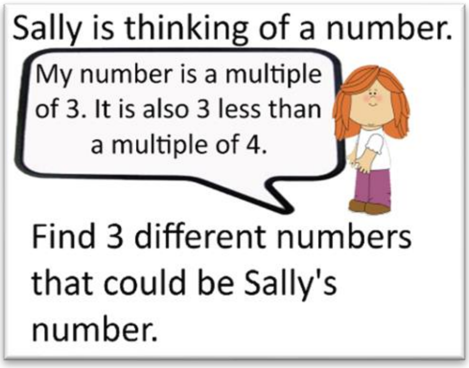
Arrange the cards to make:

a. five numbers that are multiples of 3
b. five numbers that are multiples of 7

Gold







3. Week 8 Medium-term plan and small steps

Area of Maths: Multiplication & Division	**Week: 8**		
Lesson Number	**Focus of lesson**	**Small Steps**	**Notes:**
1	Identify Factors & Multiples including finding all factor pairs. Investigate common factors.	• Make arrays from 8 tiles/counters/cubes. How many different ways? E.g. 8 × 1, 4 × 2. • Show that 8 can be made from 'multiples' of 2 (2+2+2+2) and also from 'multiples' of 4 (4+4). • Explain that 8, 1, 2 and 4 are all factors of 8. • Show that 8 is a multiple of 1, 2 and 4. (link to × tables) 8 has 4 factors. • Could also show using Cuisenaire. • Repeat for 18, 24, etc. List the factors, record the arrays. • Reveal the following factors of 48 as shown: 48 × 1 24 × 2 12 × 4 6 × 8 3 × 16 What do you notice? Discuss the double and halve strategy as a means of finding factor pairs. • Find all factor pairs for 12. Now find factor pairs for 20. Display side by side. What do you notice? Do 12 and 20 have any factors that are the same?	Necessary vocab: factor, multiple Ensure teacher modelling of mathematical language and pupil use. **https://nrich.maths.org/5158**

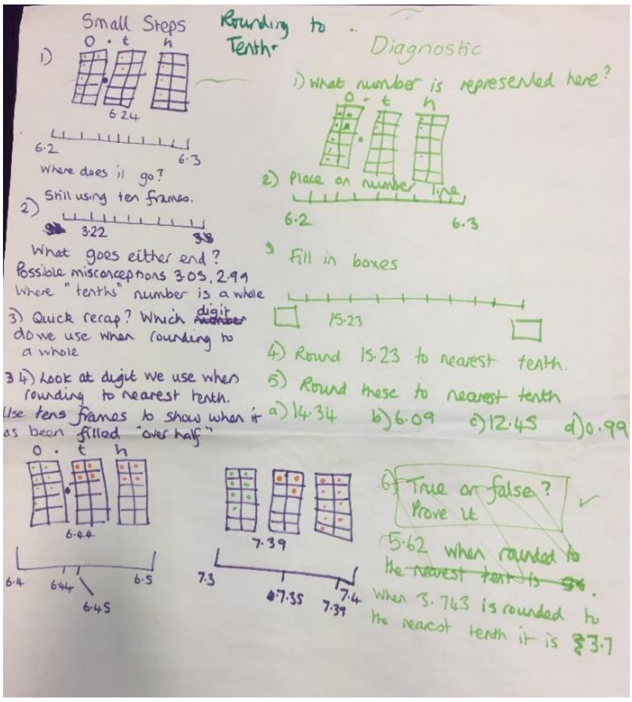
		• Reveal that 1, 2, and 4 are common factors of 12 and 20. • Repeat with 8 and 24. Record the common factors.	
2	Identify common multiples	• Use Numicon or Cuisenaire. Make a line of 3s and a line of 4s. Stop when both lines are exactly the same length. How much in each line? 12. Explain that 12 is a multiple of 3 and 12 is a multiple of 4 so we say 12 is a common multiple of 3 and 4. • Repeat with 4 and 5. 3 and 7, etc. • Move to x table list of 3 and 5. Can they identify the common multiples? Circle these. • Provide contextual problem – e.g. 2 lighthouses that flash: 1 every 3 minutes, 1 every 4 minutes. • Model how to create a table to show the common multiples (times at which both flash together). • Apply common multiples in further context problems.	Vocabulary = multiple, common multiple, common factor **https://nrich.maths.org/5158**
3	Prime numbers/ composite numbers	• Find the factors of 8 using tiles/counters/cubes to make arrays. Now try making arrays using 7. What do you notice? How many possibilities? • Could also show using Cuisenaire. • Explain that some multiples have more than 2 factors. These are called composite numbers. Explain that some factors have only 2 factors (themselves and 1); we call these prime numbers. • Show that all the prime numbers will only make ‘sticks’ when we arrange that number of cubes.	Vocabulary – composite, prime

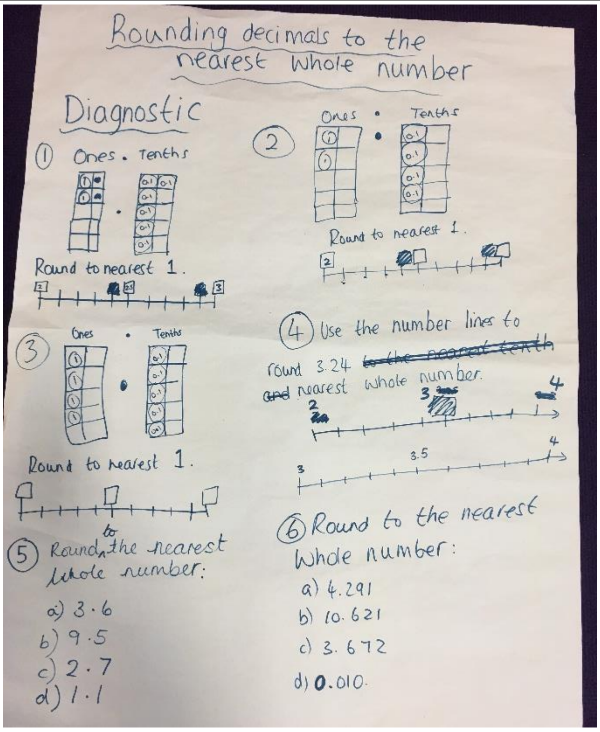
		• Investigate numbers 1–20. Sort into prime and composite. Evidence this using diagrams.	
4	Square numbers	• Provide pupils with small square tiles or cubes. Ask them to use the tiles/cubes to make the smallest square possible. How many tiles have they used? 4. Explain that 4 is a square number. We can make a square from this many square tiles. Can they make a square using 5 tiles? No. • Could also show using Cuisenaire. • Continue, challenging pupils to find all the possible squares they can make. Record in books as shown:  • Display all the square numbers pupils have found up to 100. Introduce the symbol ^{2}. Use the language e.g. $2^2 = 2 \times 2 = 4$. 4 is a square number. Record for all square numbers. • Show calculations involving square numbers, e.g. $2^2 + 3^2 = 4 + 9 = 16$. • Solve similar problems. • Look at contextual problems with square numbers, e.g. tiling a patio.	Vocabulary – square numbers, squared
5	Cube numbers	• Show children a 2 2 arrangement of cubes. Now add another 2 × 2 arrangement. Here we have 2 layers of 2 × 2. How many cubes altogether? 8. • Show children we can work this out by doing 2 × 2 × 2. 8 is a cube number. • Ask pupils to do the same for 3 × 3 and 3	Vocabulary – cube number, cubed

		layers. How many cubes? Record 3 × 3 × 3 = 27. • Can we record the cube numbers that come next without making them (as this might take too long!). • Introduce the symbol ³. Record 8 = 2 × 2 × 2 = 2³. • Repeat for other cube numbers. • Complete calculations involving cubed numbers, e.g. 2³ + 3³ = 8 + 27 = 35. • Solve similar. • Explore contextual problems for cube numbers.	

4. Examples of collaborative planning from surgeries

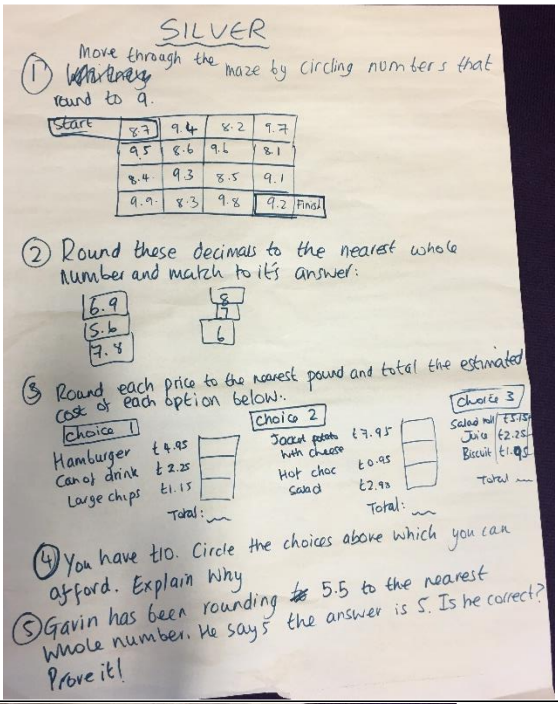
<table>
<tr>
<td>Diagnostic</td>
<td>





</td>
</tr>
</table>

Silver


SILVER

1) Move through the ~~[illegible]~~ maze by circling numbers that round to 9.

Start	8.7	9.4	8.2	9.7
	9.5	8.6	9.6	8.1
	8.4	9.3	8.5	9.1
	9.9	8.3	9.8	9.2 Finish

2) Round these decimals to the nearest whole number and match to its answer:

6.9 / 5.6 / 7.8 — 8 / 7 / 6

3) Round each price to the nearest pound and total the estimated cost of each option below.

Choice 1: Hamburger £4.95; Can of drink £2.25; Large chips £1.15; Total:

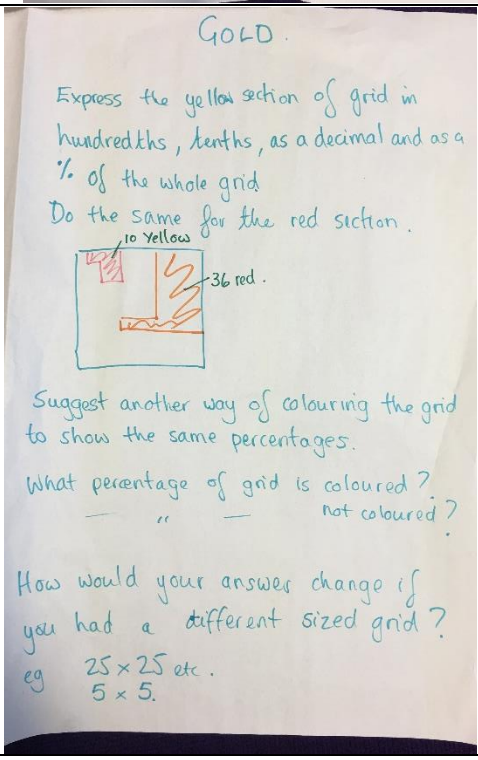
Choice 2: Jacket potato with cheese £7.95; Hot choc £0.95; Salad £2.93; Total:

Choice 3: Salad roll £3.15; Juice £2.25; Biscuit £1.95; Total

4) You have £10. Circle the choices above which you can afford. Explain why

5) Gavin has been rounding 5.5 to the nearest whole number. He says the answer is 5. Is he correct? Prove it!

Gold


GOLD.

Express the yellow section of grid in hundredths, tenths, as a decimal and as a % of the whole grid

Do the same for the red section.

10 Yellow

36 red.

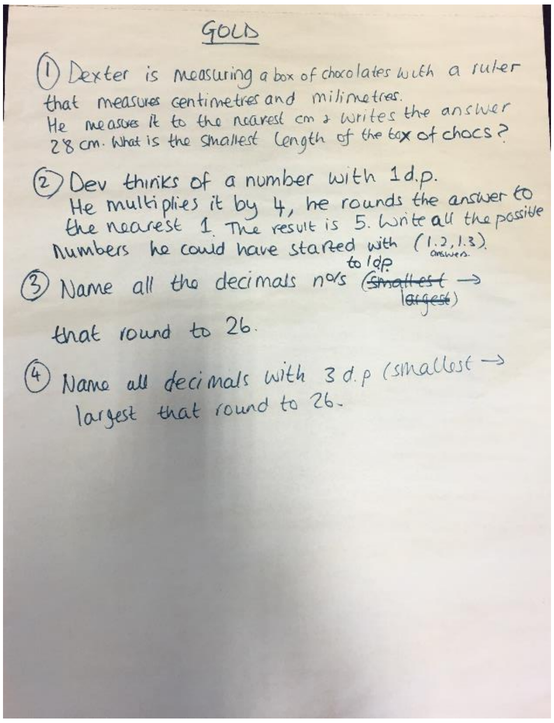
Suggest another way of colouring the grid to show the same percentages.

What percentage of grid is coloured?

— " — not coloured?

How would your answer change if you had a different sized grid?

eg 25 × 25 etc.
5 × 5.


GOLD

1) Dexter is measuring a box of chocolates with a ruler that measures centimetres and milimetres. He measures it to the nearest cm & writes the answer 28 cm. What is the smallest length of the box of chocs?

2) Dev thinks of a number with 1 d.p. He multiplies it by 4, he rounds the answer to the nearest 1. The result is 5. Write all the possible numbers he could have started with (1.2, 1.3). answers

3) Name all the decimals no's to 1dp (~~smallest~~ → ~~largest~~) that round to 26.

4) Name all decimals with 3 d.p (smallest → largest that round to 26.

Further appendix 10: 2SLS analysis of impact in the presence of non-compliance

As outlined in the impact evaluation chapter, the analysis in the presence of non-compliance was intended to be analysed using a two staged least squares (2SLS) instrumental variables regression model. However, the results of the Wu-Hausman test indicated that we should use a mixed linear regression model instead. The updated results in Version 2.0 outline these updated results. The updated results couldn't be provided in Version 1.0 of the report because of the impact of Covid-19 on access to the ONS Secure Research Service, preventing these results being updated within the publication timeline.

Table 32 Compliance analysis

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
Mathematics attainment (PTM10)	1,339 (243)	34.5 [33.7; 35.3]	1,491 (225)	34.7 [33.9; 35.4]	2,689 (1,242; 1,502)	0.02 [-0.05; 0.09]	0.418

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

https://educationendowmentfoundation.org.uk

@EducEndowFoundn

Facebook.com/EducEndowFoundn