ReflectED

Evaluation Report

December 2022

Louise Gascoine, Louise Tracey, Caroline Fairhurst, Lyn Robinson-Smith, David Torgerson, Carole Torgerson, Kerry Bell

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

We do this by:

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

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

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

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

0207 802 1653

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

The project was independently evaluated by a team from the University of York and Durham University and comprised of: Louise Gascoine, David Torgerson, Louise Tracey, Caroline Fairhurst, Lyn Robinson-Smith, Imogen Fountain, Carole Torgerson, Kerry Bell, and Louise Elliott.

The co-principal investigators were Dr Louise Gascoine and Professor David Torgerson.

Contact details:

Name: Dr Louise Gascoine

Address: School of Education,

Durham University,

Leazes Road,

Durham, DH1 1TA

United Kingdom

Email: louise.gascoine@durham.ac.uk

Acknowledgements

The evaluation team would like to acknowledge the patience of all involved with this report, including the Education Endowment Foundation, in relation to delays caused by the Covid-19 pandemic in accessing and analysing data and writing up this evaluation.

The evaluation team would like to thank all the schools, teachers, and pupils that were involved in this evaluation for giving their time and efforts in participating in the evaluation of ReflectED. Also thank you to Jenny Roche (previously of the University of York) for her contribution to the writing of the Statistical Analysis Plan and on data collation, cleaning, and analysis.

We would also like to thank the team at Rosendale Primary School and London Connected Learning Centre for their cooperation in all aspects of this evaluation.

This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Executive summary

ReflectED focuses on metacognitive skills, with the aim of assisting pupils to reflect on 'learning moments and strategies' to support attainment (ReflectED, 2021. This includes the skills of setting goals, assessing progress, and identifying personal strengths and challenges. It has been developed by Rosendale Primary School, in partnership with London Connected Learning Centre and the National Education Trust.

ReflectED is a whole-school approach delivered through a series of pre-prepared lesson plans from Reception year to Key Stage 2 (including technology aided reflection and other visual scaffolds). It consists of 28, weekly, half-hour lessons, delivered by a teacher over five school terms. Pupils are taught strategies to monitor and manage their own learning. They are supported to apply and practise strategies across the curriculum, record reflections on learning, and review this over time. Pupils have opportunities to consider, which strategies are most effective to them. Implementation of ReflectED involves a launch and training event for headteachers and lead practitioners (nominated teachers responsible for leading the interventions in their schools). Termly training sessions are held at local hubs for lead practitioners and cascaded to other teachers in their schools.

ReflectED was evaluated using a randomised controlled trial. A total of 112 schools were randomly allocated to the intervention or to continue with their 'business as usual provision'. The aim was to use KS2 tests in reading and mathematics to look at the impact of the programme on Year 5 pupils after 2 years. The process evaluation used focus groups, lesson observations, and interviews. The intervention began in February 2018 and ended in July 2019. It was preceded by an efficacy trial, which involved 24 schools and was completed in January 2016.

Table 1: Key conclusions

Key conclusions
Pupils who participated in ReflectED made the equivalent of 1 month's less progress in the primary outcome related to Key Stage (KS) 2 mathematics, on average, compared to pupils in control schools. This result has a very high security rating.
Pupils who participated in ReflectED made the equivalent of no months' progress in the primary outcome related to KS2 reading, on average, compared to pupils in other schools. This result has a very high security rating.
Pupils eligible for Free School Meals (FSM) who took part in the intervention, made the equivalent of 1 months' progress in KS2 mathematics, on average, compared to pupils in control schools. They made no months' progress in KS2 reading, on average, compared to pupils in other schools.
Teacher surveys indicate that training led by Rosendale Primary School and cascaded in ReflectED schools was well received. Staff and pupils spoke positively about the programme, and they felt that it facilitated increased pupil's awareness of their own learning. Pupils in more than half of ReflectED schools did not complete two weekly online reflections. Verbal and written reflections were completed in addition to use of the online platform.
Post-intervention teacher surveys from ReflectED schools suggested evidence of metacognition in pupil reflections. It was not possible to obtain surveys from all intervention schools, therefore this finding should be treated with caution.

EEF security rating

These findings have a very high security rating. This was an effectiveness trial, which tested whether the intervention worked under everyday conditions in a large number of schools. This was a well-designed two-arm cluster randomised trial where relatively few pupils [< 9%] who started the trial were not included the final analysis, this led to the very high security rating, for both the KS2 mathematics and KS2 reading primary outcomes

Additional findings

Pupils in ReflectED schools made, on average the equivalent of 1 months' less progress in the primary outcome of KS2 mathematics, compared to pupils in other schools. They made, on average the equivalent of no months' progress in the primary outcome of KS2 reading, compared to pupils in other schools. As with any study, there is always some uncertainty around the result. While this is our best estimate, the result is consistent with negative effects of up to 2 months less progress and positive effects of up to 1 month of additional progress for KS2 mathematics. There were also negative effects of up to 2 months less progress and positive effects of up to 2 months of additional progress for KS2 reading.

Secondary outcomes relating to grammar, punctuation, and spelling saw pupils in ReflectED schools making on average, the equivalent of 1 months' additional progress, compared to pupils in other schools. KS1 pupils who received the intervention made on average, the equivalent of no additional months' progress in the secondary outcome for raw reading and mathematics scores, compared to pupils in other schools.

Pupils in ReflectED schools who were eligible for FSM, showed that they made, on average, the equivalent of an additional 1 months' progress in KS2 mathematics compared to FSM pupils in other schools. They made, on average, the equivalent of no months' additional progress in KS2 reading.

The results of this evaluation do not align with existing literature and the EEF Toolkit, which highlight the potential value of metacognition for pupil outcomes in mathematics (e.g. Desoete and De Craene, 2019). ReflectED is a multi-component intervention that is theoretically underpinned by a concept that has been described as complex and 'fuzzy' (Wellman, 1985). There is also an argument that differing conceptualisations in literature present challenges defining, operationalising, and 'testing' metacognition (Gascoine *et al.,* 2017). Extensive evidence (EEF, 2021) highlights the value of metacognition in supporting pupils to think about their own learning. An important part of the evaluation was to actively enquire about the use of metacognition across all schools participating in the trial, as there was evidence of this approach in schools that did not deliver the intervention.

Cost

ReflectED cost each school around £861.78 per year, or £8.13 per pupil, per year when averaged over 3 years. This figure included costs for staff training and ongoing support through termly local hub sessions. Start-up costs were low, with lead staff attending a training launch and meetings hosted by Rosendale Primary School, in addition to charges for licence fees and materials.

Impact

Table 2: Summary of impact on primary outcomes

Outcome/ Group	Effect size (97.5%/95% confidence interval)	Estimated months' progress	EEF security rating	No. of pupils	P Value	EEF cost rating
KS2 Mathematics	-0.05 (97.5% CI -0.18, 0.09)	-1	🔒🔒🔒🔒🔒	4,148 (2,139; 2,009)	0.43	£ £ £ £ £
KS2 Reading	0.01 (97.5% CI -0.10, 0.11)	0	🔒🔒🔒🔒🔒	4,149 (2,133; 2,016)	0.91	£ £ £ £ £
KS2 Mathematics (FSM subgroup)	0.05 (95% CI -0.10, 0.20)	1	N/A	1,330 (702; 628)	0.51	£ £ £ £ £
KS2 Reading (FSM subgroup)	0.04 (95% CI -0.09, 0.17)	0	N/A	1,332 (700; 632)	0.52	£ £ £ £ £

Introduction

Background

ReflectED is a whole-school intervention focused on metacognition and developing metacognition to enhance primary school pupils' 'ability to think about their learning, assess their progress, set and monitor goals, identify strengths and challenges in their learning and develop a learning dialogue between pupil and teacher' (ReflectED, 2021). ReflectED was developed by Rosendale Primary School. Before this current trial began, ReflectED was being used in Rosendale Primary School, but also with a range of other schools that had previously accessed ReflectED training delivered by Rosendale Primary School and/or their associated online ReflectED materials.[1]

A previous evaluation of ReflectED was conducted in the school year 2014/2015 (Motteram, *et al.,* 2016); this was an efficacy trial funded by the Education Endowment Foundation (EEF). The efficacy trial, focusing primarily on mathematics attainment, showed a positive impact on mathematics attainment (+4 months, Hedges' g effect size 0.30, 95% confidence interval [CI] -0.04 to 0.63, p=0.08) in Year 5, though the CI for this result ranges from 0 to 8 additional months' progress. Given the positive findings arising from the efficacy trial, the EEF funded the present larger scale effectiveness trial to evaluate the impact of ReflectED when delivered as a whole-school approach, at scale and over a longer time (five school terms in total, as opposed to three school terms in the previous efficacy trial). This report presents findings of an independent evaluation of the impact of ReflectED on educational attainment in mathematics and reading (primary outcomes) and metacognition (secondary outcome). We report here the findings of a pragmatic, two-armed, cluster randomised controlled trial (RCT).

Metacognition

In the Teaching and Learning Toolkit (EEF, 2021) metacognition and self-regulation are rated very highly in terms of progress made (+7 months) versus the (very low) cost to implement interventions in this strand of the toolkit—these claims are made based on evidence that is described as extensive in the toolkit.

Metacognition refers to thinking about thinking, an individual's own knowledge of their own thinking or learning or 'knowledge and cognition about cognitive phenomena' (Flavell, 1976, p. 906). It is important to note that definitions of metacognition in the literature are broad and wide-ranging, there is a risk that defining metacognition has become 'fuzzy' (Wellman, 1985) and overly defined to a point where its core meaning becomes lost. For this reason, it is important to consider the definition of metacognition applied in the ReflectED intervention and where this sits within the broader theoretical and evaluative literature that focuses on metacognition.

The aims of ReflectED as stated above, focus on pupil's ability to think about their own learning, align with the definition of metacognition and self-regulation in the EEF's Teaching and Learning Toolkit (EEF, 2021, p. 1 para) as being something that works to help pupils to:

> *… think about their own learning more explicitly, often by teaching them specific strategies for planning, monitoring and evaluating their learning. Interventions are usually designed to give pupils a repertoire of strategies to choose from and the skills to select the most suitable strategy for a given learning task.*

The EEF guidance report on metacognition and self-regulated learning (Quigley, *et al.,* 2018, p. 8) presents a definition of self-regulation as being 'the extent to which learners […] are aware of their strengths and weaknesses and the strategies they use to learn'. They define self-regulation as the overarching concept and place within this metacognition, cognition, and motivation. Not dissimilarly to the widely cited definitions in Veenman, *et al.,* (2005), and Quigley, *et al.,* (2018) make a distinction between metacognitive knowledge (declarative knowledge about a task and the strategies available to an individual in engaging with it) and metacognitive skilfulness or skills (the procedural knowledge required to regulate learning activities). Quigley, *et al.,* (2018, pp. 8–9) define metacognition as 'the ways learners monitor and purposefully direct their learning', with metacognitive knowledge concerning knowledge of self as a learner and skills relating to the development of strategies that can be used to 'enhance their learning and improve'. The description of

[1] These online materials were redacted from general access and available only to intervention schools during the evaluation period.

what the ReflectED intervention seeks to do encompasses both metacognitive knowledge (pupils thinking about their own learning) and skills (assessing progress and setting and monitoring goals).

Previously common understandings about metacognition and its perceived lack of development in children under the age of 8 years old are now widely challenged and there is research evidence to the contrary (Bartsch, *et al.,* 2003; Gascoine, *et al.,* 2017; Kuhn, 1999; Veenman, *et al.,* 2004). The value of metacognition for pupils of all ages, including children as young as 3 and 4 years old is clear and there is a wealth of research knowledge to support this (Leutwyler, 2009; Wall, 2008; Wall, *et al.,* 2013; Whitebread, *et al.,* 2009).

The importance of the role of metacognition in relation to pedagogy has been demonstrated in research that shows positive links between metacognition and student outcomes including attainment (Akyol, *et al.,* 2010; Diginath and Büttner, 2008; Kuyper, *et al.,* 2000; Prins, *et al.,* 2006). Recently, Kyriakides, *et al.* (2020, p. 2) noted the 'prognostic' value of metacognition for supporting academic achievement and spotlighted a shift in focus from cognitive to more recognition of the value of non-cognitive outcomes. There is growing evidence in the literature that metacognition is an important predictor for attainment in core subjects including mathematics (Desoete, 2009; Desoete and De Craene, 2019; Quigley, *et al.,* 2018). More recently, although focused on an adolescent age range rather than the primary school age range of this study, Wang, *et al.* (2021) in a series of longitudinal studies, noted the value of growth mindset alongside metacognitive skills and how they are mutually supportive of each other in terms of student engagement with mathematics. Jones, *et al.* (2020) also explored explicit metacognitive strategy training concurrently alongside working memory training and found that, in a double blind RCT, the group that received a combination of working memory and metacognitive strategy training showed and sustained greater improvements.

As we have described, metacognition is a complex concept with many (sometimes competing and sometimes overlapping) definitions and applications in a pedagogical sense. Growth mindset (Dweck, 2016) is part of the ReflectED approach to metacognition, there is recent research evidence (Yeager, *et al.,* 2019), although not focused on the primary school age group of this evaluation that, supports the value of growth mindset in improving achievement.

Rationale

The previous RCT evaluation of ReflectED (Motteram, *et al.,* 2016) was an efficacy trial that focused on the impact on mathematics attainment for Year 5 pupils (aged 9–10 years) and was implemented in 30 schools across five areas (1,858 pupils in total received the intervention). The results of this previous evaluation were described as promising by the EEF and showed a positive impact on progress in mathematics (+4 months). Motteram, *et al.* (2016) commented that this finding was potentially limited by existing pedagogy in the intervention schools that already included an embedded and explicit focus on metacognitive and reflective skills. This is the main reason that supported the rationale for this larger scale effectiveness trial of ReflectED.

The results of the previous evaluation of ReflectED (Motteram, *et al.,* 2016) also showed a small, but not statistically significant, negative impact (-2 months, Hedges' g effect size -0.15, 95% CI -0.59 to 0.29, p=0.50) on reading—this informs the focus in this current evaluation on the impact of ReflectED on the primary outcomes of both mathematics and reading for KS2. It is important to consider the results of the previous trial (Motteram, *et al.,* 2016) in conjunction with the knowledge that it was a much smaller scale efficacy trial than this effectiveness trial, there was some missing data and there is also the possibility that the timetabling of ReflectED contributed to the results (e.g. if ReflectED lessons took away from literacy or related time in a school's timetable but not numeracy or mathematics for the schools involved in the intervention). This current evaluation of ReflectED differs from the earlier evaluation primarily because of its sample size (number of participating schools and pupils) and the length of the intervention period (five school terms as compared to the three school terms of the earlier evaluation). For full details of the changes relating to delivery, content, and design please see Appendix C Table 1.

Design overview

The evaluation design of this study includes an impact evaluation and an implementation and process evaluation (IPE) to explore the impact of ReflectED, as a whole-school intervention, on attainment in mathematics and reading at Key Stage (KS) 2, as well as for KS1. We also looked at the impact of ReflectED on metacognition in the KS2 cohort, using the Junior Metacognitive Awareness Inventory (JrMAI) (Sperling, *et al.,* 2002; Sperling, *et al.,* 2012) at both the beginning and end of the evaluation period when the KS2 pupils were in Year 5 and then Year 6, respectively.

The study design is a pragmatic, two-armed, cluster RCT where participating schools were randomised to receive either the ReflectED intervention or to continue with business as usual (usual teaching/control condition). The cohort that forms the primary focus of the intervention in the participating schools are KS2 pupils who were in Year 5 in January 2018 through to the end of their last year of primary school (Year 6) in summer 2019. A secondary focus was on a KS1 cohort in Year 1 as of January 2018 and in Year 2 at the end of the evaluation period.

The IPE was conducted concurrently with the impact evaluation explored over the five academic terms of the intervention period, and the experiences of schools, teachers, and pupils as they implemented or took part in ReflectED. The IPE gathered evidence of experiences using a range of data collection methods including focus groups, interviews, and observations. The integration and concurrence of the impact evaluation and IPE facilitate an evaluation design that considers the impact (on attainment) of ReflectED and at the same time, by considering implementation and process, potential reasons for the impact results.

Intervention

The intervention being tested in this evaluation is **ReflectED**, it is applied as a whole-school intervention, where the whole school from Early Years Foundation Stage (EYFS) through to Year 6 at the end of KS2 participate. In a report for the EEF, Anders, *et al.* (2017) adopted the following definition of Complex Whole-School Interventions (CWSI):

> *… as an intervention that combines multiple components that interact with one another within a context and aims to produce change (Moore et al., 2015). CWSIs may have many potential 'active ingredients' (Oakley et al., 2006). Other elements of complexity may include measurement of a range of outcomes, or targeting different levels of the organisation.*

Arguably ReflectED can be described as a CWSI as there are multiple components (e.g. ReflectED lessons, pupil reflections, shared language to talk about metacognition, and cascaded staff training) to the intervention—these are explained in more detail below. The aim is that these multiple (metacognition focused) components work together to produce change (i.e. metacognitive strategies to improve attainment), and within this that school staff work together and beyond their own schools in 'regional hubs' with other schools participating in the evaluation. What follows will describe the components of the ReflectED intervention in more detail.

Although the ReflectED intervention includes weekly ReflectED lessons for all year groups, it is intended that ReflectED (as delivered in this evaluation) becomes embedded across the whole school. The ReflectED lesson plans are not the only component of the intervention. Other key components include facilitating opportunities for pupil reflection on their learning in school. It was intended that pupils recorded reflections were 'multimedia' in their nature (including audio recordings of verbal reflections and/or images of written reflections or work that pupils have produced) and uploaded to an app called **Seesaw**. However, participating schools were also provided with a variety of examples of paper-based reflection templates (see technical notes Appendix E) and guidance about reflecting across the school. At the beginning and end of ReflectED lessons pupils are asked to attribute themselves a colour ('performance tag colours'), a way of self-assessing where their learning is in relation to a given activity. The performance tag colours were:

- Red: I don't understand anything;
- Yellow: I'm unsure and still need a little bit of help;
- Green: I understand the learning; and
- Blue: I am confident enough to coach.

Training materials provided to participating schools via the initial launch training describe the 'Key approaches to teaching ReflectED' as:

- Talk—to promote 'growth mindset', to encourage and to feedback to pupils, and to share their reflections with their peers.
- Learning from others—learning and listening as key components of reflecting on learning. The ReflectED intervention advocates for pupils working in mixed ability pairs, with regular changes of partners.
- Continuous reflective behaviour and language across the curriculum—teacher modelling of reflective behaviour and language.

- Opportunities to reflect—multiple opportunities to reflect in the week, including and beyond the specific weekly ReflectED lesson.
- Valuing and using reflections—teacher engagement with pupil reflections and providing constructive feedback on them. The importance of pupils recognising that their reflections are looked at.
- Making mistakes—promoting a classroom environment where pupils feel safe to make mistakes, and can use them to 'reflect on and improve their learning'.

Each school, upon signing a Memorandum of Understanding (MOU) to participate in the evaluation, was asked to nominate a lead practitioner for ReflectED. The lead practitioner and the headteacher were invited to attend initial training provided by the development team and to lead ReflectED in their schools—cascading the training to the remainder of their school staff. The lead practitioners also attended termly regional hub meetings and shared ReflectED-related experiences and practice with other participating schools in their geographical region.

Key components of the intervention are summarised below in Table 1, this is based on the Template for Intervention Description and Replication (TIDieR) as adapted for the EEF evaluations by Humphrey, *et al.* (2016).

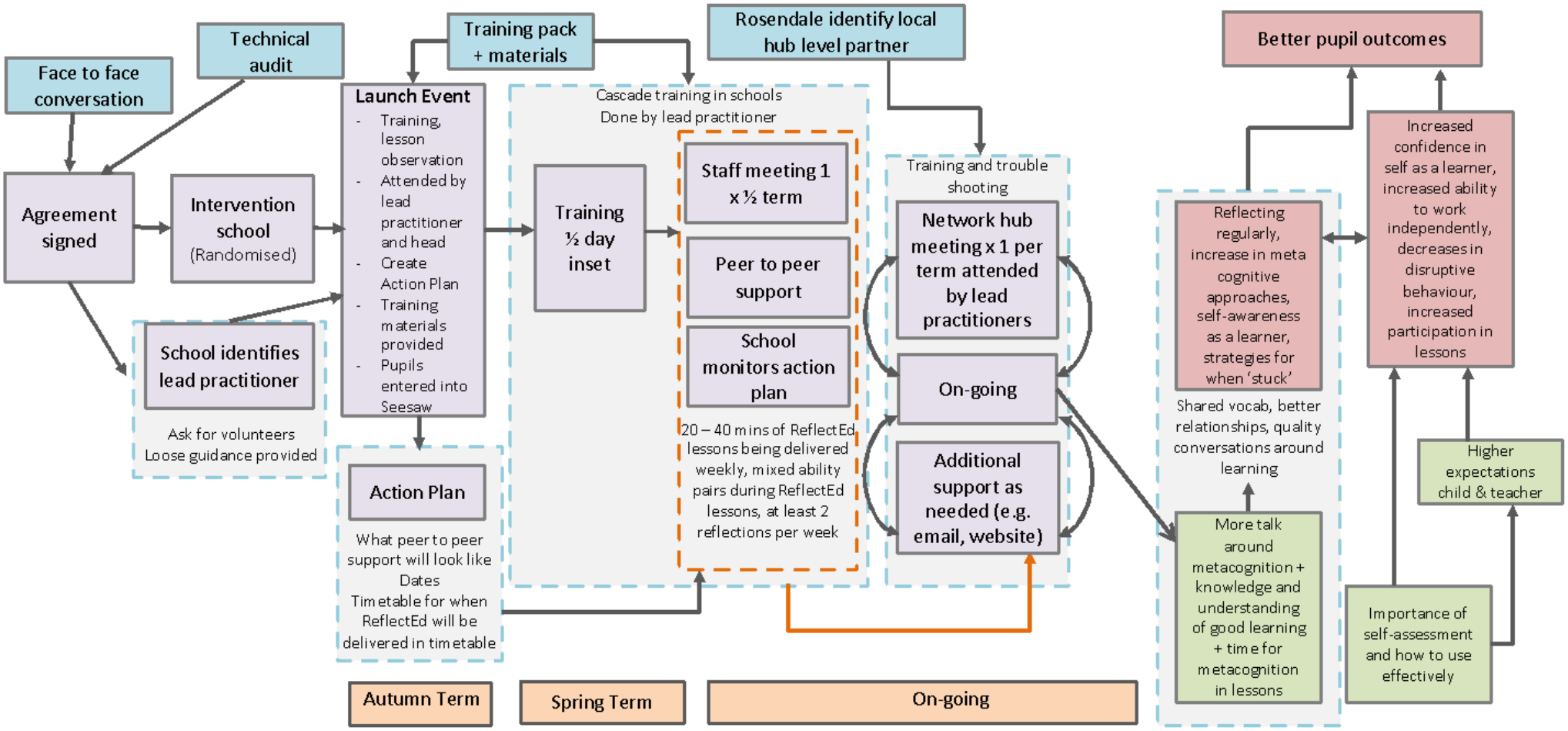
Table 1: TIDieR for ReflectED

Brief name	ReflectED
Why? Rationale, theory and/or goal of essential elements of the intervention	ReflectED focuses on teaching primary-aged students' metacognitive skills to improve attainment. Essential elements of the intervention are: talk; learning from others; technology; opportunities to reflect; reflective behaviour and language; and valuing and using pupils' reflections
Who? Recipients of the intervention	All children in intervention schools—ReflectED provides resources for students across the primary school age range (4–11 years)
What? Physical or informational materials used in the intervention	The following are provided for each school: • Lesson plans for ReflectED lessons—one lesson for each year group each week, lessons of 30–40 minutes in duration (see technical notes Appendix F) • Access to the Seesaw app for student reflections • Guidance for parents about metacognition in the classroom (this was an optional component for participating schools who were provided with a handout, which they could choose to distribute to/share with parents). No data was collected about this aspect of ReflectED • Resources for teachers (including reflection templates for 'offline' reflections if required)
What? Procedures, activities and/or processes used in the intervention	• School continuing professional development (CPD) (internal) led by lead practitioners to cascade the ReflectED training provided to them and the intervention school headteachers at the launch day • Network hub meetings led by the Rosendale Primary School project team to provide ongoing professional development and support for the lead practitioners (one per term) • ReflectED lessons delivered by school staff in intervention schools (class teachers predominantly) • Distance support from the developer
Who?	ReflectED is designed to be delivered by classroom teachers. A lead practitioner from each intervention school was responsible for disseminating ReflectED training to the remainder of their school staff, organising within school training and

Intervention providers/implementers	meetings, and attending network hub meetings termly. The lead practitioners were required to attend the initial launch training for ReflectED with the headteacher of their school. Lead practitioners were also responsible for internal quality assurance within their schools in delivering the intervention across year groups and classes
How? Mode of delivery	The delivery of ReflectED lessons is undertaken as part of regular classroom activity, by classroom teachers. ReflectED is a whole-school approach and as such, all classroom teachers should be delivering it in intervention schools
Where? Location of the intervention	Regular classrooms in participating schools that are allocated to the intervention arm of the trial. Schools will be recruited in several areas across England
When and how much? Duration and dosage of the intervention	ReflectED lessons are approximately 20–40 minutes in duration and designed to be delivered on a weekly basis. Pupils work in mixed ability pairs during ReflectED lessons. There are weekly pre-prepared lesson plans for the whole primary school-age range. The learning from these lessons needs to be applied in day-to-day teaching and learning. The evaluation period comprises five school terms of intervention with ReflectED, in schools that are not (and have not previously) been implementing ReflectED
Tailoring: adaptation of the intervention	Teachers can change the 'new skill' learned to draw on existing skills or scenarios. Cross-curricular links can be made and are encouraged. Adaptation can be made for pupil's ability level, e.g. one lesson can be taught over two separate sessions. Minimum requirement of two reflections per child per week, but no upper limit is placed on this
How well (planned)? Strategies to maximise effective implementation of the intervention	To maximise the effectiveness of the implementation the following training and CPD opportunities are provided: • Launch event (full day) attended by headteachers and lead practitioners from all intervention schools • The above includes an opportunity to visit Rosendale Primary School and see ReflectED in action • Training (1/2 day inset) delivered within each intervention school by the lead practitioner to all teachers • Network hub meetings (four in total) • Within school staff meetings focused on ReflectED (two per term) and facilitated by the lead practitioner • Quality assurance 'spot check' visits by Rosendale Primary School lead practitioners —approximately 10 schools • School action plans for each intervention school created at launch event (including what peer-to-peer support will look like, key dates, timetable for ReflectED delivery, and how to ensure that pupils are given the opportunity to reflect as part of day-to-day teaching and learning) • Peer-to-peer support • Monitoring of reflections completed by intervention school pupils on the Seesaw app • Regular, frequent email contact between ReflectED and the intervention schools • Distance telephone support if required • Technical support from the Seesaw app where required

A logic model was agreed at set up meetings and produced by the EEF (September 2018) and can be seen in Figure 1.

Figure 1: Logic model



Evaluation objectives

This evaluation sought to evaluate the impact of ReflectED, when delivered as a whole-school approach, on attainment in mathematics and reading for both KS2 (primary outcomes) and KS1 (secondary outcome). This evaluation also looked at the impact of ReflectED on metacognition as measured by the JrMAI (Sperling, *et al.,* 2002; Sperling, *et al.,* 2012) in KS2 students.

The trial protocol (Gascoine, *et al.,* 2018a) and Statistical Analysis Plan (Fairhurst and Roche, 2018) were both published by the EEF online.

The primary research questions for the impact evaluation are:

1. How effective is ReflectED in improving pupil outcomes in mathematics at the end of KS2?
2. How effective is ReflectED in improving pupil outcomes in reading at the end of KS2?

The secondary research questions for the impact evaluation are:

3. How effective is ReflectED in improving pupil outcomes in mathematics at the end of KS1?
4. How effective is ReflectED in improving pupil outcomes in reading at the end of KS1?
5. How effective is ReflectED in improving pupil outcomes in grammar, punctuation, and spelling (GPS) at the end of KS2?
6. How effective is ReflectED in improving primary outcomes for pupils with Free School Meals (FSM)?
7. Does ReflectED have an impact on metacognition for pupils in KS2, as measured by the Junior Metacognitive Awareness Inventory (JrMAI)?

The research questions for the implementation and process evaluation (IPE) are:

8. To what extent do the schools and teachers implementing ReflectED adhere to the intended model of ReflectED as a whole-school intervention?
    a. Linked to the above, how effectively has the training provided to the headteachers and the lead practitioners cascaded to the remaining teachers?
    b. What variability in the implementation of ReflectED exists across different participating settings? Are there any barriers?
9. How much of ReflectED has been delivered in the intervention schools? (e.g. how many reflections have pupils in intervention schools recorded on Seesaw? How many terms of ReflectED have been delivered?)
    a. Linked to the above, how well have different components of ReflectED been delivered and how well have participants engaged with it?
10. How does ReflectED differ from existing practice in primary schools that focuses on facilitating pupils' metacognitive skills? (Including practice in intervention schools prior to the implementation of ReflectED and activity in control schools other than ReflectED).
11. What is the reach of ReflectED across the intervention schools? For example, What proportion of training has been attended? How many reflections have been completed? Are there perceived or actual benefits for specific groups of pupils (e.g. SEN, EAL, GRT[2])?
12. Has ReflectED been adapted in any way during the intervention period? In what ways and why?
13. What evidence is there, in pupils' reflections, of metacognition and any change in this over the course of the intervention?

Ethics and trial registration

Ethical review of the evaluation of ReflectED was undertaken via the School of Education Ethics Committee at Durham University. Approval, via Chairs Action, for the initial trial protocol and first full ethics application submitted at Durham (refs 2840 and 3173 below) was granted by the ethics committee in the Department of Health Sciences at the University of York.

[2] SEN = Special Educational Needs, EAL = English as an Additional Language, GRT = Gypsy, Roma, and Traveller.

Ethical approvals were granted as follows by the School of Education Ethics Committee (Durham University) over the course of the evaluation: October 2017, Trial Protocol (ref. 2840); March 2018 addition of training surveys to IPE (ref. 3173); April 2018 Protocol Amendment (ref. 3088); April 2018 school visit documentation for IPE (ref. 3161); and June 2018 updates to documentation in relation to the General Data Protection Regulation (GDPR), including trial protocol (ref. 3166). After ethical approval was granted, two subsequent addendums were made: in June 2019 the post-intervention questionnaire (for schools); and in September 2019 the post-intervention schedule for developer interview and associated forms.

School agreement to participate in the trial was obtained in a two-step process. First, schools expressed their interest to participate in a process led by the developer team (Rosendale Primary School). Second, schools who expressed an interest were then contacted by the evaluation team about signing up for the trial, documentation was sent to prospective schools who had expressed an interest including an Memorandum of Understanding (MOU) that explained the requirements of participating in the trial, the structure of the evaluation, information about the use of data, and the responsibilities of all parties (the project team, the evaluation team, and participating schools). Headteachers were asked to give their agreement to proceed further with the involvement of their school in the evaluation and to nominate a 'lead practitioner' for the trial who would be the main point of contact and responsible for overseeing the implementation, alongside the headteacher, in their school.

The trial was registered with the International Standard Randomised Controlled Trial Number (ISRCTN) Registry before recruitment and randomisation (registration number: ISRCTN1404682).

Data protection

From September 2017, after schools had agreed to take part in the evaluation and signed MOUs had been returned, parents and carers were given the opportunity, via information sent to them via their child's school, to withdraw their child from data that was shared via the participating schools for the purpose of the evaluation of ReflectED (name, date of birth, unique pupil number [UPN], assessment data, and FSM eligibility). It was made clear in all communications to parents and carers that the ReflectED intervention was delivered at school level, so they were not being given an opportunity to withdraw their child from receiving the ReflectED intervention. We also asked that parents/carers who did not want their child to participate in an audio-recorded focus group with a Disclosure and Barring Service (DBS) checked member of the evaluation team to inform their child's teacher.

Information given to parents/carers included details of the purpose of the collection (for the evaluation of ReflectED), secure storage and transfer of data, the anonymisation of their child's data, the ethical approval of the evaluation and a statement regarding the Data Protection Act (1998). It explained that pupil-level data (as above) would be shared by schools with the evaluation team at the beginning of the evaluation period; access to pupil details was limited to specific members of the evaluation team and was required to facilitate data linkage with the National Pupil Database (NPD). This data linkage included sensitive data such as FSM status. It also explained that the NPD data would be used for the statistical analysis and will be shared in an anonymised format with the Department for Education, the EEF, Fischer Family Trust FFT Education, and potentially, in an anonymised format to the UK Data Archive.

GDPR-related addendums (from May 2018)

General Data Protection Regulation (GDPR) (EU) 2016/679 (GDPR, 2016) and the Data Protection Act (DPA, 2018) was applicable in the UK from May 2018. An addendum to the MOU was issued to participating schools in June 2018 and headteachers were asked to agree to this and circulate amended information about the use of data in the evaluation to the parents/carers of participating pupils. In this revised documentation, headteachers and parents/carers were provided with information about why we were writing to them again, a re-cap of the ReflectED intervention, a list of data being collected and why it was being collected, information about data sharing and details of the legal basis under which we were processing the data. Parents/carers were also provided with an updated list of frequently asked questions (FAQs) that included information about the data controller, data security, storage of data (including length of time), their rights in relation to the data, and what to do if they had any questions or concerns or no longer wished for their child's data to be used in the evaluation. This updated list of FAQs included a link to the University of York's Record Management and Information Governance in relation to **individual rights** University of York was named as the data controller until the data is archived at which point the EEF becomes the data controller.

In both the addendums sent to schools (MOU) and distributed by schools to individual parents/carers our legal basis for processing personal data was stated as being in line with our purpose to advance learning and knowledge, conducting this research under the GDPR on the grounds of 'public interest' i.e. processing personal data as necessary for the performance of a task (namely the research) carried out in the public interest, Article 6(1)(e),[3] and Special Category data under Article 9(2)(j)[4] of the GDPR.

All data transfers to and from schools (e.g. pupil-level data provided by schools at the beginning of the evaluation period) were sent and received via encrypted spreadsheet using the University of York's DropOff service (a secure system for file transfer). Outcome assessment data that was collected on paper (namely the JrMAI for KS2 students) were returned by recorded postal delivery and then held securely in a controlled access area in locked cabinets at York Trials Unit. The paper JrMAI forms were sent to schools in envelopes with individual pupil names on, only the participant identification (ID) was listed on the JrMAI form itself and schools were instructed to send these back without the envelopes they were delivered in, to preserve anonymity. The trial management systems and trial data are held on secure University of York servers with access limited to specified members of York Trials Unit staff. Electronic data and paper documents including identifiable personal data will be securely archived and disposed of by York Trials Unit when the youngest participating young person is aged 25. This is in line with the Limitations Act (1980) and the Records Management Code of Practice for Health and Social Care (2016). Anonymised electronic data and paper documents will be kept indefinitely. Data sharing agreements were put in place with participating schools (see technical notes Appendix G).

Copies of both the original and amended information sent to parents and carers of pupils in the participating schools can be found in the additional documentation file, and in the revised trial protocol (Gascoine, *et al.,* 2018b). All results in this evaluation report have been anonymised so that no school or individual pupil will be identifiable in this report or in any other publications that disseminate the results.

Project team

The independent evaluation was conducted by researchers from the University of York and Durham University:

- Dr Louise Gascoine, Co-Principal Investigator and Evaluation Lead, School of Education, Durham University, previously York Trials Unit, University of York (until September 2018);
- Professor David Torgerson, Co-Principal Investigator, York Trials Unit, University of York;
- Dr Louise Tracey, Co-Investigator, Department of Education, University of York;
- Caroline Fairhurst, Co-Investigator, York Trials Unit, University of York;
- Dr Lyn Robinson-Smith, (Co-Investigator), York Trials Unit, University of York;
- Professor Carole Torgerson, Co-Investigator, Department of Education, University of York, previously School of Education, Durham University (until September 2020);
- Louise Elliott, Co-Investigator, York Trials Unit, University of York;
- Imogen Fountain, Trial Support Officer), York Trials Unit, University of York; and
- Dr Kerry Bell, Co-Investigator, York Trials Unit, University of York.

The ReflectED intervention was developed and led by Rosendale Primary School and London Connected Learning Centre (CLC).

- Kate Atkins, Headteacher, Rosendale Primary School;
- Justine Paton, Lead Practitioner, ReflectED, Rosendale Primary School;
- Rachael Gallagher, Lead Practitioner, ReflectED, Rosendale Primary School;
- Julia Lawrence, Deputy Director, London CLC; and
- Marc Rowland, Head of Research School, Rosendale Primary School.

[3] (e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller General Data Protection Regulation. (2016). Art. 6 GDPR: Lawfulness of Processing. [*online*]. Available at: https://gdpr-info.eu/art-6-gdpr/ (accessed 01 November 2021).

[4] (j) processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject. Ibid.

Methods

Trial design

This evaluation of ReflectED is a pragmatic, two-armed, cluster RCT. The primary focus of the evaluation is the impact of ReflectED on pupils in KS2 (in Year 5 at the beginning of the evaluation in Autumn Term 2017 and in Year 6 at the end of the evaluation in Summer Term 2019). The secondary focus is on pupils in KS1 (Year 1 at the beginning of the evaluation in Autumn Term 2017 and in Year 2 at the end of the evaluation in Summer Term 2019). Randomisation was completed at school level; schools were randomly allocated to either the intervention (ReflectED) or control condition (teaching as usual). Table 2 summarises the trial design and is followed by information regarding the reasons for the choices made in relation to trial design.

Table 2: Trial design

Trial design, including number of arms		Pragmatic, two-arm, cluster randomised controlled trial
Unit of randomisation		School
Stratification variable		Type of school (primary, infant, junior); geographical location (North or South England); school size (<320 pupils, ≥320); and percentage of ever Free School Meals (FSM) (<20%, ≥20%)
Primary outcome (1)	Variable(s)	Mathematics attainment (Year 5 cohort)
	Measure(s) (instrument, scale, source)	Key Stage (KS) 2 mathematics score (National Pupil Database [NPD]); range 0–110
Primary outcome (2)	Variable(s)	Reading attainment (Year 5 cohort)
	Measure(s) (instrument, scale, source)	KS2 reading score (NPD); range 0–50
Secondary outcomes	Variable(s)	Grammar, punctuation, and spelling (GPS) attainment (Year 5 cohort)
	Measure(s) (instrument, scale, source)	KS2 GPS score (NPD); range 0–70
	Variable(s)	Metacognition (Year 5 cohort)
	Measure(s) (instrument, scale, source)	Junior Metacognitive Awareness Inventory (JrMAI) (collected by evaluation team); range 12–36
	Variable(s)	Mathematics attainment (Year 1 cohort)

	Measure(s) (instrument, scale, source)	KS1 mathematics score (raw data obtained directly from schools); range 0–60
	Variable(s)	Reading attainment (Year 1 cohort)
	Measure(s) (instrument, scale, source)	KS1 reading scores (raw data obtained directly from schools); range 0–40
Baseline for primary outcome 1	Variable	Mathematics attainment (Year 5 cohort)
	Measure(s) (instrument, scale, source)	KS1 mathematics score (NPD); scores 3, 9, 13, 15, 17, and 21
Baseline for primary outcome 2	Variable(s)	Reading attainment (Year 5 cohort)
	Measure(s) (instrument, scale, source)	KS1 reading score (NPD); scores 3, 9, 13, 15, 17, and 21
Baseline for secondary outcome(s)	Variable(s)	Grammar, punctuation and spelling (GPS) attainment (Year 5 cohort)
	Measure(s) (instrument, scale, source)	KS1 writing attainment point score (NPD); scores 3, 9, 13, 15, 17, and 21
	Variable(s)	Metacognition (Year 5 cohort)
	Measure(s) (instrument, scale, source)	Pre-test, - Junior Metacognitive Awareness Inventory (JrMAI) (collected by evaluation team prior to randomisation); range 12–36
	Variable(s)	Mathematics attainment (Year 1 cohort)
	Measure(s) (instrument, scale, source)	Early Years Foundation Stage Profile (EYFSP), range 1–3
	Variable(s)	Reading attainment (Year 1 cohort)
	Measure(s) (instrument, scale, source)	Early Years Foundation Stage Profile (EYFSP), range 1–3

The primary outcomes in this evaluation are academic attainment, in both mathematics and reading, for KS2 pupils who were in Year 5 (aged 9–10) in September 2017 at the time of recruitment and in Year 6 (aged 10–11) in July 2019 at the end of the evaluation period. The secondary outcomes in this evaluation were twofold: first academic attainment, in both mathematics and reading, for KS1 pupils who were in Year 1 (aged 5–6) at the time of recruitment and in Year 2 (aged 6–7) at the end of the evaluation period. Secondly, a measure of metacognition for KS2 pupils the JrMAI (Sperling,

et al., 2002; Sperling, *et al.,* 2012) . Considering debate in the field and the suitability of a self-report measure of metacognition like the JrMAI, it was deemed to not be suitable to use with KS1 students. Although there are outcome measures for both KS1 and KS2, we primarily wanted to investigate the effectiveness of ReflectED in KS2 age pupils since this was the age range of pupils involved in the efficacy trial (Motteram, *et al.,* 2016). However, since ReflectED was delivered as a whole-school intervention in this effectiveness trial, we also investigated the effectiveness in a KS1 cohort as secondary outcomes.

The control condition in this evaluation was 'teaching as usual'; schools randomised to the control condition did not receive ReflectED resources or training and they did not implement weekly ReflectED lessons. While the trial was designed with the aim of minimising burden on participating schools, inevitably all participating schools were required to complete specific tasks (beyond their usual practice) to participate in the evaluation. Intervention schools were expected to release staff to attend specific training for the ReflectED intervention; with this in mind the intervention schools were offered £1,110 per school (£185 x 6 days for training attendance) to cover the expense of the release of the lead practitioner to attend the sessions. The offer for control schools included a range of professional development opportunities (not related to ReflectED or metacognition) within the evaluation period, and a payment of £250 for each control school following the final data submission at the end of the evaluation period. Control schools were also offered free access to ReflectED (not including Seesaw licence) for the academic year after the evaluation period (2018–2019) and beyond.

Participant selection

Schools were approached in the first instance by the project team, Rosendale Primary School, supported by the evaluation team. The evaluation team attended recruitment events organised by the project team and provided support materials (e.g. information about the RCT and what this would involve) where appropriate. The project team gathered expressions of interest (EOIs) from schools that were interested in taking part in this evaluation of ReflectED. These EOIs were then shared with the York Trials Unit evaluation team and followed up by York Trials Unit sending interested schools an MOU that included further information about the evaluation and what participation would involve. Recruitment materials including the copies of the MOU and information sheets are available in the accompanying technical notes (Appendix H).

Schools were eligible to participate in the trial if they were not currently implementing ReflectED and were willing to implement the ReflectED intervention if they were randomised to the intervention group. Schools needed to be at least 1 form entry per year group in size; it was agreed that smaller schools would be considered on a case-by-case basis if any expressed an interest. All pupils in Year 1 (KS1) and Year 5 (KS2) classes at the beginning of the evaluation period (September 2017) were eligible to participate. ReflectED was delivered as a whole-school intervention in this evaluation, as such all classes (Reception year to Year 6) in the participating intervention schools were eligible to receive the intervention. The focus on Year 1 and Year 5 pupils for the gathering of data allowed an approach that minimised the data collection burden on participating schools (both control and intervention) as the outcome data collected was largely available via the NPD.

The York Trials Unit were responsible for ensuring schools had provided all of the required data before they were eligible to be randomised. Schools who chose to participate in the evaluation were only eligible for inclusion in the evaluation, and randomisation after they had:

- signed an MOU (responsibility of the headteacher in each participating school);
- provided information to the parents/carers of the pupils in Year 1 and Year 5 who were eligible for the data collection aspect of the evaluation (schools used resources provided by the evaluation team: Participant Information Sheet (PIS), opt-out consent initially, post-GDPR in 2018 this was updated to clearly specify the legal basis for processing data alongside the option to withdraw);
- provided the pre-randomisation data and information specified in the MOU:
    - Unique Pupil Numbers (UPNs), names, and date of births for participating Year 1 and Year 5 students;
    - Free School Meal (FSM) status for participating Year 1 and Year 5 students;

- Contact details for the headteacher and nominated lead practitioner;
- Completed an online pre-randomisation survey (part of the IPE and to establish baseline of usual practice prior to the intervention period); and
- Completed JrMAI (Sperling, *et al.,* 2002; Sperling, *et al.,* 2012) metacognition measures sent it back to York Trials Unit for participating Year 5 pupils.

Outcome measures

The outcome measures were listed in Table 6; what follows provides further information about the outcome measures including the baseline measures, and about how the outcome measures link to the logic model (Figure 1).

Baseline measures

All baseline measures were administered prior to randomisation. The baseline measures for the KS2 outcomes of mathematics, reading, and GPS scores were KS1 attainment point scores in mathematics, reading, and writing, respectively. These were measured in Summer Term 2015 and scores were obtained from the NPD. These are scored using the points 3, 9, 13, 15, 17, and 21; and were treated as continuous data in the analyses.

The JrMAI (see technical notes Appendix I) was administered at both baseline and outcome by teachers in the schools for pupils in KS2 and is described in more detail under the section 'Metacognition measure (JrMAI)' below.

The baseline measures for the KS1 outcomes of mathematics and reading were the EYFS Profile Early Learning Goals of G11 Mathematics – Numbers and G12 Mathematics – Shape, Space and Measures, and G09 Literacy – Reading, respectively. These were measured in Summer Term 2017 and scores were obtained from the NPD. These are scored using the points 1='emerging', 2='expected' or 3='exceeding'; and were treated as categorical data in the analyses.

Primary outcomes

We had two primary outcomes of KS2 mathematics and reading, such that the research question is formulated so that the 'success' of the intervention is defined as showing an effect on either primary outcome. These were measured in Summer Term 2019 and scores were obtained from the NPD. KS2 mathematics scores ranged from 0 to 110 and reading scores from 0 to 50.

Secondary outcomes

KS2 GPS score (range 0–70) was a secondary outcome for the Year 5 cohort. This were measured in Summer Term 2019 and scores were obtained from the NPD.

Other secondary outcomes related to the Year 1 cohort; these were KS1 raw mathematics and reading scores. These were collected for the Year 2 pupils (in Summer Term 2019) directly from participating schools, as only levels (BLW = 'Below expected standard'; PKF = 'Pre-Key stage'; WTS = 'Working towards expected standard'; EXS = 'Working at expected standard'; and GDS = 'Working at a greater depth within the expected standard') are currently reported in the NPD (e.g. for variable KS1_MATH_OUTCOME). Data were collected on a password protected and encrypted spreadsheet sent to schools by the evaluation team; when completed by the schools, this was transferred securely to the evaluation team at the University of York via DropOff, an online file transfer service.

Metacognition measure (JrMAI)

Considering the nature of metacognition and debate within the field, we deemed a self-report measure of metacognition was not suitable for KS1 students; indeed, there is not an appropriate measure in existence. The post-intervention JrMAI was administered by teachers for each pupil for baseline testing in Year 5 in Autumn Term 2017 and for outcome assessment in Year 6 in July 2019. Such a measure as the JrMAI was deemed to be suitable with KS2 students providing the additional burden for schools and pupils was carefully managed. The JrMAI has been used with students of KS2 age in several studies (Sperling, *et al.,* 2002; Sperling, *et al.,* 2012). It seeks to explore metacognitive knowledge and regulation (distinguishing between metacognitive and self-regulatory skills) and was therefore appropriate to the aims of ReflectED.

This 12-item measure, each with a 3-point Likert scale for response, should take no longer than 30 minutes to complete for each class. Version B of the instrument was used with some changes to the wording to adapt it to a UK context. Items are scored: never=1; sometimes=2; and always=3, and a total score is obtained by summing the 12 items to produce a total summary score of 12–36 with a higher score indicating greater metacognition. When up to two-item responses were missing, the missing items were completed with the mean of the completed items. Questionnaires with more than two items missing were not scored. No specific advice is provided by the developers of the JrMAI regarding the handling of missing data. This data decision was therefore made by the evaluation team to minimise the number of invalid responses, based on their experiences of common methods to manage missing item-level data in health-related patient-reported outcomes.

Sample size

Table 9 presents a summary of the sample size assumptions and the minimum detectable effect size (MDES) at protocol, randomisation, and analysis stages.

Protocol

The primary analyses compared KS2 mathematics and reading scores between the trial arms for the Year 5 cohort. We had multiple primary outcomes; hence, the research question is formulated so that the 'success' of the intervention is defined as showing an effect on either primary outcome. In this scenario, the p-value must be corrected for multiple testing and the two outcomes were tested at the 0.025 significance level. We assumed a pre- and post-test correlation of 0.65, and an intracluster correlation coefficient (ICC) between pupils of 0.13. This was based on the EEF guidance on the ICC for KS2 mathematics (0.127) and reading (0.137) scores calculated using data from the NPD 2013/2014 academic year with special schools and small schools excluded in England (EEF, 2015). We proposed to recruit 140 schools (70 intervention, 70 control) with an average of 30 pupils per class at recruitment and 15% pupil-level attrition at outcome, which would have given us 80% power to detect an effect size of 0.16 between the trial arms for either outcome.

Though the trial was not powered to detect an effect in the FSM subgroup, an analysis in the subgroup of FSM pupils was proposed as a secondary investigation. In 2016, 14.5% of primary school pupils were known to be eligible for, and claiming, FSM. Recruiting 140 schools with an average of 30 pupils per class at recruitment, we might therefore have expected approximately 610 pupils to have FSM status (obtained from NPD) in the comparison of KS2 scores. Assuming an ICC of 0.13, a pre- and post-test correlation of 0.65, alpha of 0.025, and 15% pupil-level attrition at outcome, we would have had an MDES of 0.24 with 80% power within the FSM subgroup.

At randomisation

At randomisation, with 4,526 pupils from 102 settings (average of 44.4 pupils per setting at enrolment), assuming 80% power, two-sided alpha of 0.025, pre- and post-test correlation of 0.65, an ICC of 0.13, and 15% pupil-level attrition, the MDES is estimated at 0.18 between the intervention and control groups for the primary outcomes of KS2 mathematics and reading.

The settings randomised had a mean percentage of ever FSM pupils of 25.5%; therefore, we might have expected approximately 1,154 pupils in the evaluation to be included in the FSM subgroup. With this sample size, under otherwise identical assumptions, the MDES would be 0.22.

Randomisation

Randomisation was conducted at the school level using minimisation by Caroline Fairhurst using MinimPy software (, 2013; Saghaei and Saghaei, 2011). Schools were minimised on type of school (primary, infant, junior); geographical location (North England or South England); school size (<320 pupils, ≥320 pupils); and percentage of ever FSM (<20%, ≥20%). The thresholds for school size and ever FSM were the median observed in the first 'batch' of 71 schools that were randomised. Four 'batches' of randomisations took place as and when schools fulfilled all the criteria listed above: i) 08 December 2017 of 71 schools; ii) 14 December 2017 of 9 schools; iii) 10 January 2017 of 9 schools; and iv) 17 January 2018 of 18 schools. Randomisations were performed in batches for practical reasons to allow schools to be told as early as possible after they had completed all the pre-requisites to randomisation what their allocation was so they could make plans to attend the launch training if allocated to ReflectED.

The protocol does not state that type of school would be a minimisation factor. This had to be included when it came to light that some of the participating schools were only 'juniors' (therefore no Year 1 cohort), or 'infants' (therefore no Year 5 cohort), as opposed to a full primary school containing both juniors and infants. It was therefore important to balance the different types of schools across the trial arms. This was detailed in the Statistical Analysis Plan, which was prepared at a similar time to randomisation.

Statistical analysis

Full details of the statistical analysis were published in the Statistical Analysis Plan (SAP) (Fairhurst and Roche, 2018). Analysis was conducted in Stata Version 17 (StataCorp LLC, College Station, TX, USA) using the principles of intention-to-treat, where data were available, including all schools and pupils in the groups to which they were randomised irrespective of whether or not they actually received the intervention.

Statistical significance was assessed using two-sided tests at the 2.5% level threshold for significance (alpha) for the primary outcomes, and at the 5% level for secondary outcomes. Estimates of effect are presented as an adjusted mean difference, with a 97.5% CI for the primary outcomes, a 95% CI for the secondary and other outcomes, p-values, and converted to a Hedges' g effect size.

A Consolidated Standards of Reporting Trials (CONSORT) diagram is provided to show the flow of schools and pupils through the trial, see **Figure 2**.

School and pupil characteristics and measures of prior attainment were summarised descriptively by randomised group both as randomised and as analysed in the primary analyses models. At school level, we have the following data: school type (primary, infants, juniors); description (academy, community, foundation, or voluntary school); geographical area (North England or South England); Office for Standards in Education (Ofsted) rating; setting (urban or rural); school size (in terms of pupils); percentage of pupils ever eligible for FSM; percentage of EAL pupils whose first language is not English, and percentage of pupils with SEN or Education, Health, and Care (EHC) plan. At pupil level, we have age, gender, and FSM status (NPD variable EVERFSM_6_P), as well as relevant baseline outcome measures.

Continuous measures were reported as a mean and standard deviation (SD) and categorical data as a count and percentage. The unadjusted mean difference between groups for the baseline variables for the Year 5 cohort were reported as a Hedges' g effect size with 95% CI.

For each outcome, its correlation with its associated baseline measure is reported. Histograms of outcome and (the continuous) baseline variables are presented.

Primary analysis

KS2 attainment (for both reading and mathematics, separately) for pupils in the intervention and control groups was compared using a linear mixed-model at the pupil level. Group allocation, respective baseline measure, and the minimisation factors were included as fixed effect covariates in the model. School size was dichotomised for use as a factor in the minimisation but was included as a continuous variable in the model, to take advantage of the extra information this provided. Similarly, FSM status was aggregated to the level of the school (percentage of pupils that have ever been eligible for FSM at the school in the last 6 years) for use as a factor in the minimisation but was included as a dichotomous variable at the pupil level in the model (NPD variable EVERFSM_6_P). School was included as a random effect to account for clustering by school, and robust standard errors were specified to allow for unbiased estimators under heteroscedasticity. The normality of the standardised residuals were visually checked using a QQ (Quantile – Quantile) plot, and no concerning violations of model assumptions were observed.

Secondary analysis

Secondary outcomes (KS2 raw GPS score and JrMAI for Year 5; and KS1 mathematics and reading for Year 1) were analysed in an exactly analogous way to the primary outcomes, adjusting for the appropriate associated measure of prior attainment.

Analysis in the presence of non-compliance

The ReflectED intervention was delivered across multiple levels:

- ReflectED lessons;
- school continuing professional development (CPD) (internal) led by lead practitioners; and
- network hub meetings led by the Rosendale Primary School project team to provide ongoing professional development and support for the lead practitioners.

There was a minimum requirement for each child to complete an average of two reflections per week. The evaluation period comprised five school terms of intervention with ReflectED. Data relating to number of reflections undertaken on average per pupil per week was collected from the post-intervention teacher surveys and these data were summarised.

Attendance of schools at the initial launch meeting and at each of the four hub meetings, including number attended per school, is summarised for intervention schools.

A Complier Average Causal Effect (CACE) analysis for the primary outcomes was considered to account for compliance with the intervention. An instrumental variable, two-stage least squares approach was proposed, with random group allocation as the instrumental variable (Dunn et al, 2005). Compliance was defined as a dichotomous variable at the pupil level. For each pupil in the intervention group, they were considered to have 'complied' with the intervention if:

- there was representation from their school at both the initial training (launch event) and every termly network hub meeting (initially this was planned to be five meetings, but due to capacity issues schools were only expected to attend one hub meeting across the Spring and Summer Terms 2018, so there were four meetings) (e.g. lead practitioner or other designated person attended, attendance can be by different people at each event); and
- their teacher indicated, on the post-intervention teacher survey, that for pupils in that class, on average, two reflections per pupil per week were completed. These data corresponded to the final three terms of the trial, which allowed the ReflectED intervention to have become embedded in schools.

All children in the control schools were considered to have 'complied' with their group allocation in delivering 'teaching as usual' and were not considered to have received any of the ReflectED intervention. However, the CACE analysis was ultimately not conducted due to low levels of reported compliance.

Seesaw analytics data were available for 14 intervention schools, the reasons for the low availability of Seesaw data are expanded upon in the IPE (see p. 69). In total, 11 of the schools had participating pupils in both KS1 and KS2, one was a junior school (KS2 only) and two were infant schools (KS1 only). We requested data covering the dates 1st February 2018 (initial launch event) and 23rd July 2019 (end of Summer Term 2019). Only school-level data could be extracted, and not data at a pupil, class, or year-group level. Daily aggregate summaries of activity on Seesaw was available but this related to all activity on Seesaw at the school and not just that related to metacognition, as is relevant to ReflectED. Therefore, the data were not useful for eliciting estimates of the number of reflections completed by ReflectED pupils during the evaluation, and so were not summarised or used.

Some additional, informal/non-compulsory training sessions may have taken place for the schools by the developers. It was written in the Statistical Analysis Plan that attendance by intervention schools at these would be reported but not taken account of in the CACE analysis; however, records of registers were ultimately not kept and/or made available for these.

Missing data analysis

The percentage of randomised pupils who were not included in the primary analyses models due to missing baseline or outcome data exceeded 5%. Therefore, investigations into the missing data were conducted. Multi-level logistic regression was used to predict presence or absence of outcome data, for those with valid baseline data. Predictor variables were: gender; FSM status; age; baseline score; allocation; geographical area of school; school size; type of school; description of school and school setting; Ofsted rating; and percentage of pupils at the school who were EAL, or with SEN, or EHC plan. Significant ($p<0.05$) predictors are discussed in this report. In addition, multiple imputation by chained equations was conducted to predict baseline and outcome data for the primary outcomes using: gender; FSM status; age; allocation; geographical area of school; school size; type of school; description of school; school setting; Ofsted rating; and percentage of pupils at the school who were EAL, or with SEN, or EHC plan, all of which were

available for all randomised pupils. KS2 outcome data (mathematics and reading scores) were predicted by linear regression; truncated linear regression (with lower and upper limits of 3 and 21) were used for KS1 scores, since this was the range of scores observed for pupils with non-missing data. A 'burn-in' of 10 was used, which means that the first 10 iterations of the imputation are not used to allow the iterations to converge to a stationary distribution, and 20 imputed datasets were created. (Note, that the Statistical Analysis Plan specified a burn-in of 100 and 200 imputations, but this would have taken a considerable length of time to complete, and such large values are not necessary; therefore, 10 and 20, respectively were used). The primary analyses models were then rerun within the imputed datasets and Rubin's rules used to combine the estimates.

Subgroup analyses

Pupil-level FSM status was obtained via the NPD (variable EVERFSM_6_P). The effect of the intervention on pupils who had ever been eligible for FSM was assessed via the inclusion of FSM status and an interaction term between FSM status and treatment allocation in the primary analysis models.

The primary analyses were also repeated in the subgroup of EVERFSM_6_P pupils.

Estimation of effect sizes

Effect sizes were calculated by dividing the adjusted mean difference between the intervention and control group (controlling for prior attainment and the minimisation factors) by the pooled unconditional SD obtained from a model run without these covariates:

$$ES = (\bar{Y}_I - \bar{Y}_C)^{adjusted} / s*$$

where, $(\bar{Y}_I - \bar{Y}_C)^{adjusted}$ denotes the adjusted mean difference between trial groups from the multi-level analysis model; and $s*$ denotes the pooled, unconditional SD of the two groups (square root of the sum of the between- and within-cluster variance). A CI for the effect size was calculated by dividing the confidence limits for the adjusted mean difference by $s*$. All parameters used in these calculations are provided.

Estimation of ICC

The intracluster correlation coefficient (ICC) associated with school for the pre- and post-test outcomes are provided with 95% CIs. The ICC for outcome scores was obtained from the adjusted analysis models. The ICC for baseline measures was obtained from a mixed effects linear model (or mixed effects ordinal logistic model for the EYFS Profile baseline variables, since these were considered as categorical data) with the variable of interest as the dependent variable, and accounting only for school as a random effect.

Implementation and process evaluation

The IPE was designed to be delivered concurrently with the impact evaluation, at trial set up stage and across the five school terms of ReflectED that were delivered in intervention schools between February 2018 and July 2019. Pre-intervention, the IPE aimed to establish a picture of usual practice in relation to the outcomes (mathematics, reading, and metacognition) and canvas existing knowledge and understanding of metacognition in the participating schools. In line with the EEF guidance (Humphrey, *et al.,* 2016), the IPE seeks to explore relationships between the delivery and impact on pupil outcomes of the ReflectED intervention.

Research methods

The IPE is both cross-sectional and longitudinal in design, gathering data at specified time points at baseline (pre-randomisation), during and after the intervention period. The IPE used different methods to gather data pertaining to the research questions posed and triangulated it to explore the relationship between the delivery of the intervention and its impact on student attainment. The IPE made use of data that was already being collected (by the development team) including registers of attendance at training, where appropriate, to minimise the burden of data collection on participating schools. School visits, where data was collected from intervention schools, were planned in order that different data sources were collected in the same visit (e.g. classroom observations, pupil focus groups, and teacher interviews were conducted in one visit where scheduling allowed).

Compliance was addressed by data collected by the developers, including registers of attendance at training and metrics from the Seesaw app, as well as self-reported data in the online surveys for intervention schools. Fidelity to the intervention, as set out by the developers, was explored through structured observations, self-report from teachers in the online surveys, and metrics from the Seesaw app. Data was gathered about usual practice, for both intervention and control schools, via the baseline survey. Further data about usual practice, in relation to practice within the evaluation period itself and other interventions that schools (intervention and control) may have implemented during this time, was gathered in the post-intervention online survey for teachers and headteachers. As linked to the outcome measures (primary and secondary) in the impact evaluation, data gathered about usual practice focused on usual practice in relation to mathematics, reading, and metacognition.

The data collection methods used to address the IPE research questions are summarised in Table 5, which includes details of who collected the data and how many participants there were for each data collection method. As detailed in Table 5, the approach to minimising bias and ensuring rigour in the data collection process in the IPE included:

- The baseline survey was required to be completed by all schools prior to randomisation, as a condition of the MOU.

- The structured observations of ReflectED lessons (n=16) were completed by two core members of the evaluation team (LG and LT). LG and LT developed the observation protocol with input from the ReflectED development team. Inter-observer reliability checks were conducted for quality assurance. LG and LT completed a total of four observations together initially, to ensure an acceptable level agreement using the observation protocol. CT quality assured observations (three in total) for LG (x 2) and LT (x 1). Descriptive data was agreed, and summaries were developed by both observers of joint observation visits.

- The observation data collected in the IPE was collected across the intervention period of the evaluation (June 2018 and July 2018; November 2018 and December 2018; and March 2018, April 2018, June 2018). More observation visits (n=10) were conducted with a focus on KS2 (n=6 with a KS1 focus) as KS2 outcomes are the primary outcomes in the impact evaluation.

- Launch training (n=1) and hub meeting observations (n=5) were completed by LG and LT, they attended the launch training together, and one of the hub meetings. LG and LT both attended ReflectED training delivered by the development team in June 2017, prior to the evaluation beginning in order that they both developed a good working knowledge of the ReflectED intervention. Descriptive summaries of training observations conducted jointly were agreed by both observers and the initial joint visits meant both researchers had a shared understanding of what they were looking for in the observations.

- Semi-structured interviews (classroom teachers, development team) and focus groups (pupils) were conducted with a pre-agreed semi-structured schedule, this allowed consistency in the questions asked at interview and space for additional relevant material. When interviews were conducted all participants were offered an opportunity to have prior sight of the transcribed interview and to make amendments or add comments if they felt something was not accurate.

Analysis

As can be seen in Table 5, the approach to IPE data analysis was pragmatic; different approaches to analysis were utilised for different data sources and with the relevant IPE research questions in mind. The data collection tools are available in the technical notes (Appendices J and K). The analysis conducted on the data collected in the IPE was reflective of the logic model (Figure 1), seeking to build a rich picture of the delivery of ReflectED—a complex, whole-school intervention. The logic model includes outcomes (under the heading of ‘better pupil outcomes’) that are not explicitly captured in the outcomes of the impact evaluation but are explored in the IPE data collection and analysis. For example, questions in the post-intervention online survey and in the teacher interviews and pupil focus groups address aspects of the logic model including vocabulary and pupils talk about their own learning, pupils’ confidence in their learning and participation in ReflectED lessons (and beyond), pupil understanding of metacognition, and knowledge and understanding of ‘good learning’.

Interviews and focus groups were audio-recorded, transcribed, and then data managed using NVivo Version 12 (QSR International Pty Ltd, Victoria, Melbourne, 2018). An inductive approach to data analysis was primarily used for interviews, focus groups, and free-text answers to questions in the online survey. This inductive approach was guided by the evaluation objectives; while these established the areas to be investigated, they did not influence expectation(s) about specific findings. This approach allowed us to: i) 'to condense extensive and varied raw text data into a brief, summary format; and ii) to establish clear links between the research objectives and the summary findings derived from the raw data' (Thomas, 2006, p. 237). On a practical level, this involved multiple readings of the raw data to identify frequent and dominant themes from which a coding framework was developed and the transcripts subsequently coded. Schools who participated in the interviews/focus groups were allocated a unique ID, and interview participants described only as 'teacher' or 'pupil' to assure anonymity and confidentiality.

The free-text survey responses provided in the baseline survey response to questions about what metacognition looks like (in school and classroom, headteacher and classroom teacher, respectively) were analysed differently in a process of reflexive thematic analysis (RTA) (Braun and Clarke, 2021) using a 'six-phase process for data engagement, coding and theme development' (p. 4):

1. Familiarisation with the dataset;
2. Systematic coding;
3. Generating initial themes;
4. Developing and reviewing themes;
5. Refining, defining, and naming themes; and
6. Writing up.

Given the context of this thematic analysis within a large-scale RCT of a metacognition focused intervention (ReflectED) it would have been remiss not to acknowledge the knowledge of the authors of conceptualisations of metacognition that we knew were in line with or alike those of the ReflectED intervention (e.g. growth mindset as per the ReflectED training we had observed) and that are prevalent in existing research. This was a deductive lens through which the data was viewed and interpreted. However, this was not the only lens—a reflexive approach facilitated an engagement with the data analysis that explored patterns in the data that both did and did not fit with these existing theoretical lenses. Thinking reflexively about the positionality of the authors in relation to the thematic analysis that follows on from this, the authors are the evaluators, so they are in a position of having knowledge of metacognition both separate to and linked to the ReflectED evaluation. One of the principal investigators (LG) has a specific interest in metacognition and has studied it in-depth and written about it (Gascoine, 2016; Gascoine, *et al.,* 2017)—this author led the RTA and it was important for them to engage in regular, reflective discussion, with co-authors while engaged in the analysis.

Table 3: Implementation and process evaluation (IPE) methods overview

Research methods	Data collected	Why was the data collected?	Participants or data source	Control/ intervention	Expected number collected	Time point(s)	Data analysis	Research questions addressed	Who collected the data?
Survey	Pre- and post-intervention surveys devised by YTU with input from the developers	To establish baseline and usual practice in all participating schools with a focus on practice with relevance to the outcome measures Post-intervention to contribute to understanding of compliance, fidelity, and embeddedness of ReflectED	Headteachers	Control and intervention	1 in each school	Baseline (prior to randomisation) Post-intervention	Descriptive statistics for the closed questions and those with a Likert scale Inductive and deductive coding techniques were applied depending on the question type	8 (a and b), 9 (and 9a), 10, 12	The data was collected online, using Qualtrics (an online survey platform). The survey was developed and distributed by YTU. Completion was monitored via YTU to ensure eligibility for randomisation
			Lead practitioners (if not nominated KS1 or KS2 teacher below)	Control (pre- only) and intervention	1 in each school as nominated by the school	Baseline (prior to randomisation) Post-intervention			
			KS1 teacher	Control and intervention	1 in each school as nominated by the school	Baseline (prior to randomisation) Post-intervention			
			KS2 teacher	Control and intervention	1 in each school as nominated by the school	Baseline (prior to randomisation) Post-intervention			
	Training survey	To establish key information (role and previous ReflectED training) for those who attended the training and to establish how they found it	Lead practitioners and headteachers who attended the ReflectED training	Intervention	2 from each school allocated to the intervention condition	After the initial (launch) training delivered by the development team	Descriptive statistics and inductive coding of any free-text answers given	8 (and 8a)	The data was collected online, using Qualtrics (an online survey platform). The survey was developed and distributed by YTU
Secondary data collection	Training and hub attendance registers; Seesaw analytics data	To contribute to understanding of compliance and fidelity	Intervention schools	Intervention	All intervention schools	Data for each training event and regional hub; ongoing throughout the evaluation period – data from Seesaw	CACE analysis	9, 11	This data was already collected by the development team, either via registers that they kept of attendance at training or via the analytics from the Seesaw app

Research methods	Data collected	Why was the data collected?	Participants or data source	Control/ intervention	Expected number collected	Time point(s)	Data analysis	Research questions addressed	Who collected the data?
Structured observations	Observations of ReflectED training (launch training and regional hub meetings)	To develop evaluation team knowledge and awareness of what ReflectED is expected to look like in practice and to understand how this communicated to intervention schools	The developer team delivering the training and regional hub meetings, the teachers from the intervention schools participating	Intervention	Launch training observation x 1; regional hub meeting observations x 5	Launch training at the beginning of the intervention period, regional hub meetings to be observed at time intervals spread out across the evaluation period	Description of observational data collected, using observation field notes and training resources provided to schools at the training sessions	8 (8a, 8b), 12 (hub meetings)	Members of the evaluation team, namely LG and LT, collected the observation data. Quality assurance (joint visits) were carried out with LG and LT, and LG and CT, and LT and CT
	Structured observations of ReflectED lessons in intervention schools. Observation framework agreed with developers	ReflectED lesson observations to explore what ReflectED looks like as it is delivered in a sample of the intervention schools. Focus included: fidelity; compliance; adaptations; barriers; and engagement	Academic year 2017/18: Year 5 and Year 1 teacher and class Academic year 2018/2019: Year 6 and Year 2 teacher and class	Intervention	20 school visits were planned in total	The school visits and associated observations were planned to be distributed across the evaluation period. They were spread as evenly as possible across the Summer Term 2018, Autumn Term 2018, and Spring and Summer Terms 2019	Description of observational data collected and triangulation with other data sources	8 (8a, 8b), 9a, 12	
Semi-structured interviews	Classroom teachers, headteachers	Exploring teacher experiences of ReflectED and their delivery of it. Looking at how embedded ReflectED is, adaptations and barriers (actual or perceived), perceived student engagement and the perceived value of ReflectED for pupils	Academic year 2017/18: Year 5 and Year 1 teacher and class Academic year 2018/2019: Year 6 and Year 2 teacher and class	Intervention	20 school visits were planned in total, with teacher interviews planned to be conducted in each	The school visits and associated teacher interviews were planned to be distributed across the evaluation period. They were spread as evenly as possible across the Summer Term 2018, Autumn Term 2018, and Spring and Summer Terms 2019	Inductive and deductive coding completed by LR-S and LG Some thematic analysis of free-text questions about what metacognition looks like	8 (8a, 8b), 9 (9a), 10, 12, 13	Members of the evaluation team, namely LG and LT, collected the teacher interview data. Quality assurance (joint visits) were carried out with LG and LT, and LG and CT, and LT and CT
	Developer interview (pre-specified	To develop understanding of how the intervention period has been for	Members of the developer team	N/A	Minimum 2	The developer interview was conducted after the end of the evaluation		8, 9, 12	LG and LT (evaluation team) conducted the developer

Research methods	Data collected	Why was the data collected?	Participants or data source	Control/ intervention	Expected number collected	Time point(s)	Data analysis	Research questions addressed	Who collected the data?
	interview schedule)	the developers in terms of any changes or adaptations to the intervention, and future plans				period (8 October 2019)			interview which was recorded and transcribed
Pupil focus groups	Focus groups	To explore student understanding and perceptions of ReflectED, to consider the acceptability of the ReflectED intervention for pupils	Academic year 2017/18: Year 5 and Year 1 pupils Academic year 2018/2019: Year 6 and Year 2 pupils	Intervention	4–5 pupils per focus group (20 schools expected in total)	The school visits and associated pupil focus groups were planned to be distributed across the evaluation period. They were spread as evenly as possible across the Summer Term 2018, Autumn Term 2018, and Spring and Summer Terms 2019	Inductive and deductive coding completed by LR-S and LG	8 (8b), 9a, 10, 12	Members of the evaluation team, namely LG and LT, carried out the pupil focus groups. Quality assurance (joint visits) were carried out with LG and LT, and LG and CT, and LT and CT

Costs

Programme costs

Data on costs was collected directly through correspondence between the EEF, evaluation team, and implementation team as well as through school headteachers via the post-implementation survey. In the context of the trial, some costs were subsidised by the EEF. In the Costs section of this report, we detail the costs associated with implementing the programme within a 'real-world' context. In some cases, we were required to make assumptions around costs, where this is the case, these are outlined in the Costs section.

Time costs

Data regarding the time involved at the school level to facilitate the programme were collected directly from school staff via the post-implementation survey.

Overall costs

Per pupil costs were determined by summing the total annual costs and dividing by the number of pupils we would anticipate would receive ReflectED based on a whole-school approach, assuming a school size equal to the mean school size of schools as randomised in this trial (n=318).

Timeline

Table 4: Evaluation timeline

Dates	Activity	Staff responsible/leading
Summer Term 2017	Trial set up, including the development of documentation and attendance at set up meetings	Documentation: evaluation team Set up meetings: all (evaluation team, development team – Rosendale Primary School, the EEF)
Summer Term 2017 – Autumn Term 2017	Gathering expressions of interest for potential participating schools	Development team, with support from evaluation team where necessary
Summer Term 2017 – Autumn Term 2017	School sign up (via MOU) – main trial recruitment period	Evaluation team and development team
Autumn Term 2017	Information provided to parents and carers via participating schools regarding consent and option to withdraw from data sharing	Evaluation team
October 2017 – February 2018	Baseline surveys (online) distributed to participating schools and completion monitored	Evaluation team
Autumn Term 2017 – January 2018	Rolling randomisation of schools in batches as MOUs returned	Evaluation team
Between 11 January 2018 and 8 February 2018 (five iterations)	ReflectED launch training for intervention schools	Rosendale Primary School, evaluation team observing
February 2018*	Beginning of intervention period (five school terms in total)	Development team
April 2018 – May 2018	Online survey for those from intervention schools who attended the launch training (training survey)	Evaluation team
From May 2018 – end of evaluation period	Termly regional hubs	Development team, evaluation team observing some (as per IPE methods section)
From May 2018 – June 2019	School visits for IPE data collection incorporating structured lesson	Evaluation team

Dates	Activity	Staff responsible/leading
	observations, teacher interviews, and pupil focus groups	
May 2019 – July 2019	Outcome data collection from schools (KS2 JrMAI and KS1 secondary outcomes)	Evaluation team, YTU statisticians
July 2019	Intervention period ends	N/A
July 2019	Post-intervention online survey (control and intervention schools)	Evaluation team
August 2019 – November 2019	Complete IPE data analysis and write up	Evaluation team
March 2022	Access NPD data	Evaluation team – YTU statisticians
April 2022	Submission of draft report to the EEF	Evaluation team
December 2022	Submission of final report to the EEF	Evaluation team

* The initial anticipated start date of the intervention period was January 2018; this was pushed back to February 2018 to accommodate the last iteration of the launch training (February 2018) and owing to the recruitment taking longer than anticipated.

Impact evaluation

Participant flow including losses and exclusions

A total of 205 settings were assessed for eligibility (Figure 2). Of these, 93 (45.4%) were ineligible and 112 (54.6%) were randomised. The 112 settings consisted of 94 primary schools, 8 junior schools, and 10 infant schools. In total, 10 of these settings were randomised in pairs (n=5 pairs), such that both settings in the pair were allocated to the same group; four of these pairs consisted of linked junior and infant settings, and the other pair consisted of two primary schools that were in the same federation. Therefore, the number randomised was 107: 54 to the intervention group; and 53 to the control group. This equates to 56 in each group for the full 112 settings, within which there were a total of 9,116 participating pupils: 4,602 intervention; 4,514 control.

Year 1 (KS1)

In total, 104 of the 112 settings had pupils in Year 1 (primary and infant schools only): 52 intervention; 52 control. There were a total of 4,590 participating Year 1 pupils in these schools: 2,321 intervention; 2,269 control.

Year 2 (KS2)

In total, 102 of the 112 settings had pupils in Year 5 (primary and junior schools only): 51 intervention; 51 control. There were a total of 4,526 participating Year 5 pupils in these schools: 2,281 intervention; 2,245 control.

Three schools (all control, total of 226 pupils) formally withdrew before post-testing. NPD data were requested for all 8,890 pupils in the remaining 109 settings (4,602 intervention, 4,288 control): Year 1 total 4,457 (2,321 intervention, 2,136 control); and Year 5 total 4,433 (2,281 intervention, 2,152 control).

Valid KS1 and KS2 mathematics scores were available for 4,148 pupils (91.6% of the 4,526 randomised: intervention n=2,139, 93.8%; control n=2,009, 89.5%), and these were included in the analysis for the mathematics primary outcome. Valid KS1 and KS2 reading scores were available for 4,149 pupils (91.7% of the 4526 randomised: intervention n=2,133, 93.5%; control n=2,016, 89.8%), and these were included in the analysis for the reading primary outcome.

Table 7 shows the MDES at each stage of the study. At analysis, for the KS2 mathematics outcome, we had an attrition rate of 8.4%; therefore, 4,148 pupils were included in the analysis (from 99 settings, average of 41.9 pupils per setting). The correlation between KS1 and KS2 mathematics scores was observed to be 0.70 and the ICC for KS2 mathematics associated with setting was 0.14. Therefore, the estimated MDES, with 80% power and alpha of 0.025, was 0.18. A total of 1,330 pupils (from 94 schools, 14.1 per school) had ‘ever FSM’ status and valid KS1 and KS2 mathematics data; within this subgroup, the correlation between KS1 and KS2 mathematics scores was 0.66 and the ICC was 0.15. Therefore, the MDES for KS2 mathematics in the FSM subgroup (alpha 0.025) was 0.22. At analysis, for the KS2 reading outcome, we had an attrition rate of 8.3%; therefore, 4,149 pupils were included in the analysis (from 99 settings, average of 41.9 pupils per setting). The correlation between KS1 and KS2 reading scores was observed to be 0.67 and the ICC for KS2 reading associated with setting was 0.08. Therefore, the estimated MDES, with 80% power and alpha of 0.025, was 0.15. A total of 1,332 pupils (from 94 schools, 14.2 per school) had ‘ever FSM’ status and valid KS1 and KS2 reading data; within this subgroup, the correlation between KS1 and KS2 reading scores was 0.64 and the ICC was 0.09. Therefore, the MDES for KS2 reading in the FSM subgroup (alpha 0.025) was 0.19.

Figure 2: Participant flow diagram (CONSORT).

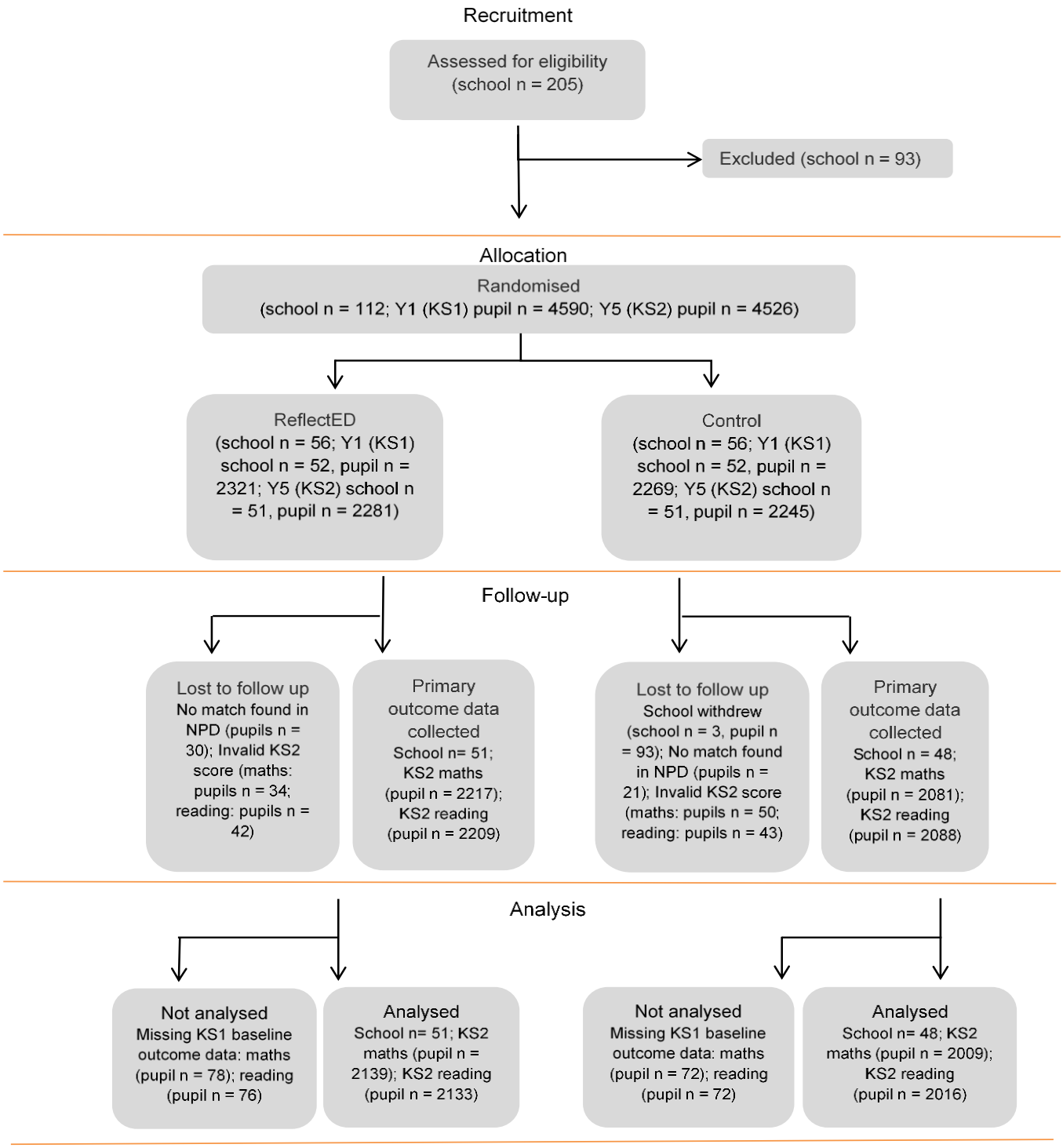


Table 5: Minimum detectable effect size (MDES) at different stages

		Protocol		Randomisation[a]		Analysis (mathematics)		Analysis (reading)	
		Overall	FSM	Overall	FSM	Overall	FSM	Overall	FSM
MDES		0.16[b]	0.24[b]	0.18[b]	0.22[b]	0.18	0.22	0.15	0.19
Pre-test/ post-test correlations	Level 1 (pupil)	0.65	0.65	0.65	0.65	0.70	0.66	0.67	0.64
	Level 2 (class)	–	–	–	–	–	–	–	–
	Level 3 (school)	–	–	–	–	–	–	–	–
Intracluster correlations (ICCs)	Level 2 (class)	–	–	–	–	–	–	–	–
	Level 3 (school)	0.13	0.13	0.13	0.13	0.14	0.15	0.08	0.09
Alpha		0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025
Power		0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		Two	Two	Two	Two	Two	Two	Two	Two
Average cluster size		30	4	44.4	11.3[c]	41.9	14.1	41.9	14.2
Number of schools	Intervention	70	70	51	51	51	49	51	49
	Control	70	70	51	51	48	45	48	45
	Total	140	140	102[d]	102[d]	99	94	99	94[e]
Number of pupils	Intervention	2,100	305	2,281	582	2,139	702	2,133	700
	Control	2,100	305	2,245	572	2,009	628	2,016	632
	Total	4,200	610	4,526	1,154	4,148	1,330	4,149	1,332

[a] The number of schools and pupils listed in the randomisation columns represent the confirmed number of eligible pupils in participating schools; the figures differ slightly from those listed in the published Statistical Analysis Plan, since these were based on the best estimate at the time of writing the Statistical Analysis Plan. [b] Allowing for 15% pupil-level attrition. [c] The settings randomised had a mean percentage of ever FSM pupils of 25.5%. [d] A total of 102 schools had pupils in Year 5 at randomisation who could potentially have contributed to the KS2 maths and reading outcome analysis. [e] In five schools, none of the pupils for whom valid pre- and post-test KS2 data were available were ever FSM; hence this analysis only included pupils from 94 schools.

Attrition

Data could be matched in the NPD for 4,382 pupils (96.8%; intervention n=2,251, 98.7%; control n=2,131, 94.9%).

The measure of prior attainment for the primary outcome of KS2 raw mathematics score was KS1 mathematics attainment point score. This was available for 4,221 pupils (93.3%; intervention n=2,166, 95.0%; control n=2,055, 91.5%). KS2 mathematics score was available for 4,298 pupils (95.0%; intervention n=2,217, 97.2%; control n=2,081, 92.7%). Pupils were included in the primary analysis model for mathematics score if they had data for the outcome (KS2 mathematics score) and covariates (KS1 mathematics score, ever FSM status, and the school-level factors of type of school, geographical location, and school size). No data were missing for the covariates of ever FSM status, and the school-level factors of type of school, geographical location, and school size; however, of the 4,298 with valid outcome data, 150 (3.5%) were missing KS1 mathematics data. This led to the inclusion of 4,148 pupils (91.6% of the 4,526 randomised: intervention n=2,139, 93.8%; control n=2,009, 89.5%). Across treatment groups the ratios (analysed to randomised) were 2,139:2,281 in the intervention arm and 2,009:2,245 for the control arm (Table 8).

The measure of prior attainment for the primary outcome of KS2 raw reading score was KS1 reading attainment point score. This was available for 4,220 pupils (93.2%; intervention n=2,165, 94.9%; control n=2,055, 91.5%). KS2 reading score was available for 4,297 pupils (94.9%; intervention n=2,209, 96.8%; control n=2,088, 93.0%). Pupils were included in the primary analysis model for reading score if they had data for the outcome (KS2 reading score) and covariates (KS1 reading score, ever FSM status, and the school-level factors of type of school, geographical location, and school size). No data were missing for the covariates of ever FSM status, and the school-level factors of type of school, geographical location, and school size; however, of the 4,297 with valid outcome data, 148 (3.4%) were missing KS1 reading data. This led to the inclusion of 4,149 pupils (91.7% of the 4,526 randomised: intervention n=2,133, 93.5%; control n=2,016, 89.8%). Across treatment groups the ratios (analysed to randomised) were 2,133:2,281 in the intervention arm and 2,016:2,245 for the control arm (Table 8).

Table 6: Pupil-level attrition from the trial (primary outcomes)

KS2 mathematics		Intervention	Control	Total
Number of pupils	Randomised	2,281	2,245	4,526
	Analysed	2,139	2,009	4,148
Pupil attrition (from randomisation to analysis)	Number	142	236	378
	Percentage	6.2	10.5	8.4
KS2 reading				
Number of pupils	Randomised	2,281	2,245	4,526
	Analysed	2,133	2,016	4,149
Pupil attrition (from randomisation to analysis)	Number	148	229	377
	Percentage	6.5	10.2	8.3

Pupil and school characteristics

Characteristics of the 107 settings as randomised are presented by group in Table 9. Approximately 60% of the recruited schools were located in the North of England. The median school size was 320 pupils, and a quarter of pupils in each school, on average, had ‘ever FSM’ status. The characteristics were well balanced across groups.

Table 10 provides the school level and pupil level characteristics for the KS1 population as randomised. Most of the schools were academies (n=42, 40.4%), and 37 (35.6%) were community schools; the remaining quarter were foundation or voluntary schools (n=25, 24.0%). Around 70% (n=73) were in urban areas, and 74 (71.2%) had a current Ofsted rating of 'Good' or 'Outstanding' (as opposed to 'Requires improvement', 'Inadequate,' or 'No Ofsted rating available'). In the intervention arm, 49.7% of pupils were male and 23.2% had 'ever FSM' status, which was similar to the control arm (50.2% and 22.0%, respectively). The distribution of scores for the three EYFS Profile variables considered were similar across the two groups. In general, the characteristics were well balanced between the groups, except that a higher proportion of schools in the intervention group (n=41, 78.8%) than the control group (n=33, 63.5%) had a current Ofsted rating of 'Good' or 'Outstanding'.

Table 11 provides the school level and pupil level characteristics for the KS2 population as randomised. Most of the schools were academies (n=39, 38.2%), and 37 (36.3%) were community schools; the remaining quarter were foundation or voluntary schools (n=26, 25.5%). Around 73 (71.6%) were in urban areas, and 70 (68.6%) had a current Ofsted rating of 'Good' or 'Outstanding' (as opposed to 'Requires improvement', 'Inadequate,' or 'No Ofsted rating available'). In the intervention arm, 51.5% of pupils were male and 33.0% had 'ever FSM' status, which was similar to the control arm (51.9% and 31.7%, respectively). For the pre-test JrMAI and KS1 mathematics, reading, and writing outcomes, scores were slightly higher in the control group, than the intervention group. However, in general, the characteristics were well balanced between the groups, except that a higher proportion of schools in the intervention group (n=39, 76.5%) than the control group (n=31, 60.8%) had a current Ofsted rating of 'Good' or 'Outstanding'.

Table 12 provides the school level and pupil level characteristics for the KS2 population as analysed in the primary analysis models.

Table 7: School-level baseline characteristics of groups as randomised (n=107 randomised units)

School level (categorical)		Intervention group		Control group	
		n/N (missing)	Count (%)	n/N (missing)	Count (%)
School type	Infant or Junior	54/54 (0)	5 (9.3)	53/53 (0)	5 (9.4)
	Primary		49 (90.7)		48 (90.6)
Geographical area (England)	North	54/54 (0)	33 (61.1)	53/53 (0)	32 (60.4)
	South		21 (38.9)		21 (39.6)
Percentage FSM	<20	54/54 (0)	25 (46.3)	53/53 (0)	25 (47.2)
	≥20		29 (53.7)		28 (52.8)
School size	<320	54/54 (0)	30 (55.6)	53/53 (0)	29 (54.7)
	≥320		24 (44.4)		24 (45.3)
School level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)
Percentage FSM		54/54 (0)	24.7 (16.8)	53/53 (0)	25.1 (17.2)
School size		54/54 (0)	323.6 (165.2)	53/53 (0)	318.7 (169.0)

Table 8: School-level and pupil-level baseline characteristics of groups as randomised for the KS1 (Year 1) trial population (n=104 settings, n=4,590 pupils)

School level (categorical)		Intervention group		Control group	
		n/N (missing)	Count (%)	n/N (missing)	Count (%)
School type	Infant	52/52 (0)	5 (9.6)	52/52 (0)	5 (9.6)
	Primary		47 (90.4)		47 (90.4)
Description	Academy	52/52 (0)	17 (32.7)	52/52 (0)	20 (38.5)
	Community school		24 (46.2)		18 (34.6)
	Foundation or voluntary school		11 (21.2)		14 (26.9)
Geographical area (England)	North	52/52 (0)	32 (61.5)	52/52 (0)	32 (61.5)
	South		20 (38.5)		20 (38.5)
School setting	Rural	52/52 (0)	15 (28.8)	52/52 (0)	16 (30.8)
	Urban		37 (71.2)		36 (69.2)
Ofsted rating	Outstanding	52/52 (0)	9 (17.3)	52/52 (0)	7 (13.5)
	Good		32 (61.5)		26 (50.0)
	Requires improvement or Inadequate or No Ofsted rating available		11 (21.2)		19 (36.5)
School level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)
School size		52/52 (0)	315.1 (160.0)	52/52 (0)	294.3 (151.7)
Percentage of SEN pupils with a statement or EHC plan		52/52 (0)	1.5 (1.5)	52/52 (0)	1.6 (1.7)
Percentage of eligible pupils with SEN support		52/52 (0)	13.1 (6.0)	52/52 (0)	12.3 (5.8)
Percentage of EAL pupils		52/52 (0)	15.8 (23.0)	52/52 (0)	14.5 (22.9)
Percentage FSM		52/52 (0)	24.7 (17.0)	52/52 (0)	25.9 (17.7)
Pupil level (categorical)		n/N (missing)	Count (%)	n/N (missing)	Count (%)
Gender	Male	2,321/2,321 (0)	1153 (49.7)	2,269/2,269 (0)	1,139 (50.2)
FSM	Eligible	2,321/2,321 (0)	538 (23.2)	2,269/2,269 (0)	500 (22.0)
EYFS Profile Mathematics – Numbers (NPD variable FSP_MAT_G11)	1 (‘emerging’)	2,250/2,321 (71)	510 (22.7)	2,092/2,269 (177)	437 (20.9)
	2 (‘expected’)		1,405 (62.4)		1,349 (64.5)
	3 (‘exceeding’)		335 (14.9)		306 (14.6)
EYFS Profile Mathematics – Shape,	1 (‘emerging’)	2,250/2,321 (71)	473 (21.0)	2,092/2,269 (177)	380 (18.2)

space, and measures (NPD variable FSP_MAT_G12)	2 (‘expected’)		1,446 (64.3)		1,433 (68.5)
	3 (‘exceeding’)		331 (14.7)		279 (13.3)
EYFS Profile Literacy – Reading (NPD variable FSP_LIT_G09)	1 (‘emerging’)	2,250/2,321 (71)	552 (24.5)	2,092/2,269 (177)	456 (21.8)
	2 (‘expected’)		1279 (56.8)		1,255 (60.0)
	3 (‘exceeding’)		419 (18.6)		381 (18.2)
Pupil level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)
Age (years)		2,321/2,321 (0)	5.5 (0.3)	2,269/2,269 (0)	5.5 (0.3)

Table 9: School-level and pupil-level baseline characteristics of groups as randomised for the KS2 (Year 5) trial population (n=102 settings, n =4,526 pupils)

School level (categorical)		Intervention group		Control group	
		n/N (missing)	Count (%)	n/N (missing)	Count (%)
School type	Junior	51/51 (0)	4 (7.8)	51/51 (0)	4 (7.8)
	Primary		47 (92.2)		47 (92.2)
Description	Academy	51/51 (0)	18 (35.3)	51/51 (0)	21 (41.2)
	Community school		21 (41.2)		16 (31.4)
	Foundation or voluntary school		12 (23.5)		14 (27.5)
Geographical area (England)	North	51/51 (0)	30 (58.8)	51/51 (0)	30 (58.8)
	South		21 (41.2)		21 (41.2)
School setting	Rural	51/51 (0)	14 (27.5)	51/51 (0)	15 (29.4)
	Urban		37 (72.5)		36 (70.6)
Ofsted rating	Outstanding	51/51 (0)	7 (13.7)	51/51 (0)	4 (7.8)
	Good		32 (62.7)		27 (52.9)
	Requires improvement or Inadequate or No Ofsted rating available		12 (23.5)		20 (39.2)
School level (continuous)		n/N(missing)	Mean (SD)	n/N(missing)	Mean (SD)
School size		51/51 (0)	320.7 (159.7)	51/51 (0)	305.1 (154.7)
Percentage of SEN pupils with a statement or EHC plan		51/51 (0)	1.6 (1.5)	51/51 (0)	1.7 (1.7)
Percentage of eligible pupils with SEN support		51/51 (0)	13.2 (5.6)	51/51 (0)	12.5 (5.8)
Percentage of EAL pupils with English not as first language		51/51 (0)	15.7 (23.2)	51/51 (0)	16.4 (25.4)

Percentage FSM		51/51 (0)	25.4 (16.4)	51/51 (0)	26.5 (17.6)	
Pupil level (categorical)		n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Gender	Male	2,281/2,281 (0)	1,175 (51.5)	2,245/2,245 (0)	1,165 (51.9)	
FSM	Eligible	2,281/2,281 (0)	752 (33.0)	2,245/2,245 (0)	711 (31.7)	
Pupil level (continuous)		n/N(missing)	Mean (SD)	n/N(missing)	Mean (SD)	Hedges' g effect size (95% CI)
Age (years)		2,281/2,281 (0)	9.5 (0.3)	2,245/2,245 (0)	9.5 (0.3)	N/A
Pre-test JrMAI		2,184/2,281 (97)	27.6 (3.3)	2,107/2,245 (138)	27.8 (3.3)	-0.06 (-0.12, 0.00)
KS1_MATPOINTS		2,166/2,281 (115)	16.3 (3.5)	2,055/2,245 (190)	16.5 (3.5)	-0.06 (-0.12, 0.00)
KS1_READPOINTS		2,165/2,281 (116)	16.4 (3.9)	2,055/2,245 (190)	16.6 (3.7)	-0.05 (-0.12, -0.01)
KS1_WRITPOINTS		2,165/2,281 (116)	15.2 (3.8)	2,055/2,245 (190)	15.5 (3.7)	-0.06 (-0.12, 0.00)

Table 10: School-level and pupil-level baseline characteristics of groups as analysed for the KS2 (Year 5) mathematics outcome (n=99 settings; n =4,148 pupils) and reading outcome (n=99 settings; n =4,149 pupils)

School level (categorical)		Intervention group		Control group	
		n/N (missing)	Count (%)	n/N (missing)	Count (%)
School type	Junior	51/51 (0)	4 (7.8)	48/48 (0)	4 (8.3)
	Primary		47 (92.2)		44 (91.7)
Description	Academy	51/51 (0)	18 (35.3)	48/48 (0)	20 (41.7)
	Community school		21 (41.2)		14 (29.2)
	Foundation or voluntary school		12 (23.5)		14 (29.2)
Geographical area (England)	North	51/51 (0)	30 (58.8)	48/48 (0)	28 (58.3)
	South		21 (41.2)		20 (41.7)
School setting	Rural	51/51 (0)	14 (27.5)	48/48 (0)	15 (31.3)
	Urban		37 (72.5)		33 (68.8)
Ofsted rating	Outstanding	51/51 (0)	7 (13.7)	48/48 (0)	4 (8.3)
	Good		32 (62.7)		26 (54.2)
	Requires improvement or Inadequate or No Ofsted rating available		12 (23.5)		18 (37.5)
School level (continuous)		n/N(missing)	Mean (SD)	n/N(missing)	Mean (SD)
School size		51/51 (0)	320.7 (159.7)	48/48 (0)	307.6 (158.2)

Percentage of SEN pupils with a statement or EHC plan		51/51 (0)	1.6 (1.5)	48/48 (0)	1.7 (1.8)	
Percentage of eligible pupils with SEN support		51/51 (0)	13.2 (5.6)	48/48 (0)	12.1 (5.6)	
Percentage of EAL pupils with English not as first language		51/51 (0)	15.7 (23.2)	48/48 (0)	16.2 (25.8)	
Percentage FSM		51/51 (0)	25.4 (16.4)	48/48 (0)	25.5 (17.4)	
KS2 Mathematics						
Pupil level (categorical)		n/N(missing)	Count (%)	n/N(missing)	Count (%)	
Gender	Male	2,139/2,139 (0)	1,110 (51.9)	2,009/2,009 (0)	1,038 (51.7)	
FSM	Eligible	2,139/2,139 (0)	702 (32.8)	2,009/2,009 (0)	628 (31.3)	
Pupil level (continuous)		n/N(missing)	Mean (SD)	n/N(missing)	Mean (SD)	Hedges' g effect size (95% CI)
Age (years)		2,139/2,139 (0)	9.5 (0.3)	2,009/2,009 (0)	9.5 (0.3)	N/A
Pre-test JrMAI		2,062/2,139 (77)	27.6 (3.3)	1,900/2,009 (109)	27.8 (3.2)	-0.07 (-0.13, -0.01)
KS1_MATPOINTS		2,139/2,139 (0)	16.3 (3.5)	2,009/2,009 (0)	16.6 (3.4)	-0.08 (-0.14, -0.02)
KS1_READPOINTS		>2,129/2,139 (<10)	16.5 (3.9)	2,009/2,009 (0)	16.7 (3.6)	-0.08 (-0.14, -0.02)
KS1_WRITPOINTS		>2,129/2,139 (<10)	15.3 (3.8)	2,009/2,009 (0)	15.6 (3.5)	-0.09 (-0.15, -0.02)
KS2 Reading						
Pupil level (categorical)		n/N(missing)	Count (%)	n/N(missing)	Count (%)	
Gender	Male	2,133/2,133 (0)	1,108 (51.9)	2,016 /2,016 (0)	1,041 (51.6)	
FSM	Eligible	2,133/2,133 (0)	700 (32.8)	2,016 /2,016 (0)	632 (31.3)	
Pupil level (continuous)		n/N(missing)	Mean (SD)	n/N(missing)	Mean (SD)	Hedges' g effect size (95% CI)
Age (years)		2,133/2,133 (0)	9.5 (0.3)	2,016/2,016 (0)	9.5 (0.3)	N/A
Pre-test JrMAI		2,056/2,133 (77)	27.6 (3.3)	1,908/2,016 (108)	27.8 (3.2)	-0.06 (-0.12, 0.00)
KS1_MATPOINTS		2,133/2,133 (0)	16.4 (3.5)	2,016/2,016 (0)	16.6 (3.4)	-0.07 (-0.13, -0.01)
KS1_READPOINTS		2,133/2,133 (0)	16.5 (3.8)	2,016/2,016 (0)	16.7 (3.6)	-0.07 (-0.13, -0.01)
KS1_WRITPOINTS		2,133/2,133 (0)	15.3 (3.8)	2,016/2,016 (0)	15.6 (3.5)	-0.08 (-0.14, -0.01)

Outcomes and analysis

Primary analysis

KS2 mathematics

Of the 4,526 randomised pupils from the 102 KS2 settings, a valid KS1 mathematics point score was obtained from 4,221 pupils (93.3%; intervention n=2,166, 95.0%; control n=2,055, 91.5%), from 99 schools (51 intervention, 48 control). The ICC associated with school for KS1 mathematics score was 0.09 (95% CI 0.06 to 0.12). In total, a valid KS2 mathematics score was available for 4,298 pupils (95.0%; intervention n=2,217, 97.2%; control n=2,081, 92.7%), from 99 schools (51 intervention, 48 control). A mean of 74.8 (95% CI 73.8 to 75.9) was observed in the intervention arm and 76.5 (95% CI 75.4 to 77.5) in the control arm. Figure 3 presents histograms for the KS1 and KS2 mathematics scores.

Valid KS1 and KS2 mathematics scores were available for 4,148 pupils (91.6% of the 4,526 randomised: intervention n=2,139, 93.8%; control n=2,009, 89.5%), and these were included in the analysis for the mathematics primary outcome. The correlation between the KS1 and KS2 mathematics scores was 0.70. The unadjusted mean difference is -1.62 (95% CI -3.07 to -0.17, p = 0.03; Appendix D Table 2). The adjusted mean difference in KS2 mathematics score between the intervention and control groups was -1.14 (97.5% CI –4.36 to 2.08, Appendix Table 2). The estimated Hedges' g effect size was -0.05 (97.5% CI -0.18 to 0.09), which relates to approximately 1 month's less progress in the intervention group (Table 13). No statistically significant difference was observed (p = 0.43). The total variance used to calculate the effect size was 588.69; the sum of 547.78 (random variation between pupils, within-cluster variance) and 40.92 (heterogeneity between schools, between-cluster variance). The ICC for KS2 mathematics score associated with school was 0.14 (95% CI 0.1 to 0.17).

In the unadjusted analysis the difference in the primary mathematics outcome was statistically significant (i.e. p = 0.03); however, in the adjusted analysis, which corrects for a small chance imbalance in baseline KS1 mathematics score, this difference is smaller and no longer statistically significant.

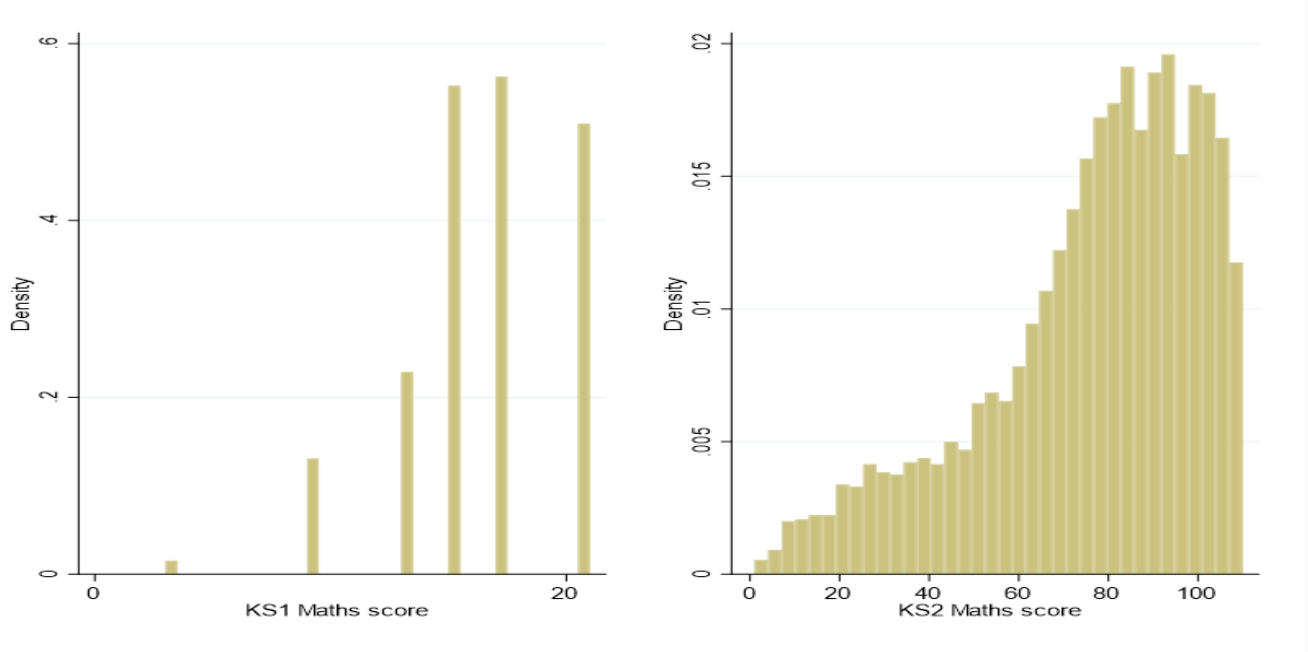


Figure 3: Histogram of KS1 and KS2 maths scores. KS = Key Stage

KS2 reading

Of the 4,526 randomised pupils from the 102 KS2 settings, a valid KS1 reading point score was obtained from 4,220 pupils (93.2%; intervention n=2,165, 94.9%; control n=2,055, 91.5%), from 99 schools (51 intervention, 48 control). The ICC associated with school for KS1 reading score is 0.08 (95% CI 0.05 to 0.11). In total, a valid KS2 reading score was available for 4,297 pupils (94.9%; intervention n=2,209, 96.8%; control n=2,088, 93.0%), from 99 schools (51 intervention, 48 control). A mean of 32.4 (95% CI 31.9 to 32.8) was observed in the intervention arm and 32.5 (95% CI 32.1 to 32.9) in the control arm. Figure 4 presents histograms for the KS1 and KS2 mathematics scores. Valid KS1 and KS2 reading scores were available for 4,149 pupils (91.7% of the 4,526 randomised: intervention n=2,133, 93.5%; control n=2,016, 89.8%), and these were included in the analysis for the reading primary outcome. The correlation between the KS1 and KS2 reading scores was 0.67. The unadjusted mean difference is -0.12 (95% CI -0.73 to 0.49, p = 0.69; Appendix D Table 2). The adjusted mean difference in KS2 reading score between the intervention and control groups was 0.06 (97.5% CI –1.06 to 1.17, Appendix D Table 2). The estimated Hedges' g effect size was 0.01 (97.5% CI -0.10 to 0.11), which relates to 0 months' additional progress in the intervention group (Table 13). No statistically significant difference was observed (p = 0.91). The total variance used to calculate the effect size was 104.66; the sum of 97.85 (random variation between pupils, within-cluster variance) and 6.81 (heterogeneity between schools, between-cluster variance). The ICC for KS2 reading score associated with school was 0.08 (95% CI 0.05 to 0.11).

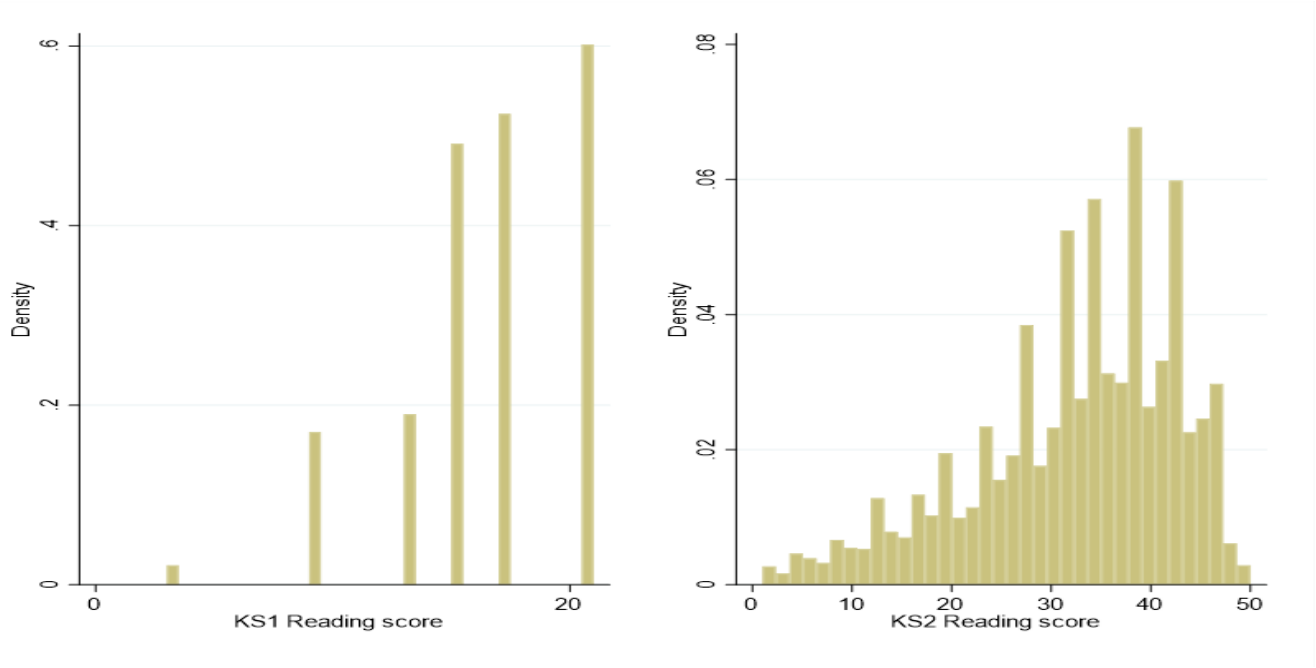


Figure 4: Histogram of KS1 and KS2 reading scores.

Secondary analysis

KS2 GPS

Of the 4,526 randomised pupils from the 102 KS2 settings, a valid KS1 writing point score was obtained from 4,220 pupils (93.2%; intervention n=2,165, 94.9%; control n=2,055, 91.5%), from 99 schools (51 intervention, 48 control). The ICC associated with school for KS1 writing score was 0.09 (95% CI 0.06 to 0.12). In total, a valid KS2 GPS score was available for 4,297 pupils (94.9%; intervention n=2,210, 96.9%; control n=2,087, 93.0%), from 99 schools (51 intervention, 48 control). A mean of 45.3 (95% CI 44.7 to 45.9) was observed in the intervention arm and 46.5 (95% CI 45.9 to 47.1) in the control arm. Technical Appendix L Figure 1 presents histograms for the KS1 writing and KS2 GPS scores. Valid KS1 writing and KS2 GPS scores were available for 4,149 pupils (91.7% of the 4,526 randomised: intervention n=2,134, 93.6%; control n=2,015, 89.8%), and these were included in the adjusted analysis. The correlation between the KS1 writing and KS2 GPS scores was 0.70. The unadjusted mean difference is -1.20 (95% CI -2.06 to -0.34, p = 0.01; Appendix D Table 2). The adjusted mean difference in KS2 GPS score between the intervention and control groups was -1.00 (95% CI -2.56 to 0.56, Appendix D Table 2). The estimated Hedges' g effect size was -0.07 (95% CI -0.18 to 0.04), which relates to approximately 1 month's less progress in the intervention group (Table 13). No statistically significant difference was observed (p = 0.21). The total variance used to calculate the effect size was 209.08; the sum of 194 (random variation between pupils, within-cluster variance) and 15.08 (heterogeneity between schools, between-cluster variance). The ICC for KS2 GPS score associated with school was 0.13 (95% CI 0.09 to 0.17).

JrMAI

Of the 4,526 randomised pupils from the 102 KS2 settings, a valid pre-test JrMAI score was obtained from 4,291 pupils (94.8%; intervention n=2,184, 95.7%; control n=2,107, 93.9%), from 101 schools (50 intervention, 51 control). The ICC associated with school for the pre-test score was 0.06 (95% CI 0.04 to 0.09). In total, a valid post-test JrMAI score was available for 3,700 pupils (81.7%; intervention n=1,942, 85.1%; control n=1,758, 78.3%), from 85 schools (44 intervention, 41 control). A mean of 27.1 (95% CI 27 to 27.3) was observed in the intervention arm and 27.2 (95% CI 27.1 to 27.4) in the control arm. Technical Appendix L Figure 2 presents histograms for the pre- and post-test scores. Valid pre- and post-test JrMAI scores were available for 3,543 pupils (78.3% of the 4,526 randomised: intervention n=1,875, 82.2%; control n=1,668, 74.3%), and these were included in the adjusted analysis. The correlation between the pre- and post-test JrMAI scores was 0.36. The unadjusted mean difference is -0.11 (95% CI -0.31 to 0.09, p = 0.29; Appendix D Table 2). The adjusted mean difference in post-test JrMAI score between the intervention and control groups was -0.13 (95% CI -0.49 to 0.22, Appendix D Table 2). The estimated Hedges' g effect size was -0.04 (95% CI -0.16 to 0.07) (Table 13). No statistically significant difference was observed (p = 0.46). The total variance used to calculate the effect size was 9.15; the sum of 8.46 (random variation between pupils, within-cluster variance) and 0.69 (heterogeneity

between schools, between-cluster variance). The ICC for post-test JrMAI score associated with school was 0.06 (95% CI 0.04 to 0.09).

KS1 mathematics

The ELG FSP_MAT_G11 correlated more highly with outcome (0.60) than FSP_MAT_G12 (0.55), and so this was used as the measure of prior attainment in the analysis model. Of the 4,590 randomised pupils from the 104 KS1 settings, a valid EYFS Profile mathematics score was obtained from 4,342 pupils (94.6%; intervention n=2,250, 96.9%; control n=2,092, 92.2%), from 100 schools (51 intervention, 49 control). The ICC associated with school for FSP_MAT_G11 was 0.08 (95% CI 0.05 to 0.11). In total, a valid KS1 mathematics score was available for 3,750 pupils (81.7%; intervention n=1,909, 82.2%; control n=1,841, 81.1%), from 91 schools (46 intervention, 45 control). A mean of 38.6 (95% CI 37.9 to 39.3) was observed in the intervention arm and 39.1 (95% CI 38.4 to 39.7) in the control arm. Technical Appendix L Figure 3 presents a histogram for the KS1 mathematics score. Valid EYFS Profile and KS1 mathematics scores were available for 3,713 pupils (80.9% of the 4,526 randomised: intervention n=1,887, 81.3%; control n=1,826, 80.5%), and these were included in the adjusted analysis. The unadjusted mean difference is -0.47 (95% CI -1.44 to 0.50, p = 0.34; Appendix D Table 2). The adjusted mean difference in KS1 mathematics score between the intervention and control groups was 0-.16 (95% CI 0-1.98 to 01.66, Appendix D Table 2). The estimated Hedges' g effect size was -0.01 (95% CI -0.13 to 0.11), which relates to approximately 0 months' additional progress in the intervention group (Table 13). No statistically significant difference was observed (p = 0.86). The total variance used to calculate the effect size was 225.44; the sum of 204.97 (random variation between pupils, within-cluster variance) and 20.47 (heterogeneity between schools, between-cluster variance). The ICC for KS1 mathematics score associated with school was 0.11 (95% CI 0.08 to 0.15).

KS1 reading

Of the 4,590 randomised pupils from the 104 KS1 settings, a valid EYFS Profile reading score was obtained from 4,342 pupils (94.6%; intervention n=2,250, 96.9%; control n=2,092, 92.2%), from 100 schools (51 intervention, 49 control). The ICC associated with school for FSP_LIT_G09 was 0.06 (95% CI 0.04 to 0.08). In total, a valid KS1 reading score was available for 3,647 pupils (79.5%; intervention n=1,896, 81.7%; control n=1,751, 77.2%), from 90 schools (46 intervention, 44 control). A mean of 26.6 (95% CI 26.1 to 27.1) was observed in the intervention arm and 26.8 (95% CI 26.3 to 27.3) in the control arm. Technical Appendix L Figure 4 presents a histogram for the KS1 reading score. Valid EYFS Profile and KS1 reading scores were available for 3,612 pupils (78.7% of the 4,526 randomised: intervention n=1,875, 80.8%; control n=1,737, 76.6%), and these were included in the adjusted analysis. The correlation between the EYFS Profile and KS1 reading scores was 0.60. The unadjusted mean difference is -0.15 (95% CI -0.87 to 0.57, p = 0.69; Appendix D Table 2). The adjusted mean difference in KS1 reading score between the intervention and control groups was 0.35 (95% CI -0.87 to 1.57, Appendix D Table 2). The estimated Hedges' g effect size was 0.03 (95% CI -0.08 to 0.14), which relates to 0 months' additional progress in the intervention group (Table 13). No statistically significant difference was observed (p = 0.57). The total variance used to calculate the effect size was 120.54; the sum of 111.92 (random variation between pupils, within-cluster variance) and 8.63 (heterogeneity between schools, between-cluster variance). The ICC for KS1 reading score associated with school was 0.08 (95% CI 0.06 to 0.11).

Analysis in the presence of non-compliance

Summaries of compliance (in relation to attendance at the training events and number of reflections completed per child) for all schools is provided in the IPE section of this report. This section focuses on summaries relating to schools (n=51) and pupils in the KS2 cohort only, since this was the population on which the primary analysis was conducted.

Data on the number of reflections completed, on average, per pupil per week, was provided on the post-intervention teacher survey. A response to this survey was received from at least one teacher in 48 (94.1%) of the 51 intervention schools with KS2 pupils, of which 44 provided a response to the question ‘How many reflections were completed, on average, per pupil per week?’ The mean response was 1.8 (SD 1.6; median 1; range 0.5 to 10). In total, 22 (50.0%) schools responded that their KS2 pupils completed at least two reflections, on average, per week.

Among the 51 intervention schools with KS2 pupils, 50 (98.0%) sent at least one staff member to the initial training, 42 (82.4%) to a hub meeting in Spring/Summer Terms 2018, 32 (62.7%) to a hub meeting in Autumn Term 2019, 36 (70.6%) to a hub meeting in Spring Term 2019 and 28 (54.9%) to a hub meeting in Summer Term 2019. One of the schools did not attend any of the meetings, two (3.9%) attended one meeting in total, five (9.8%) attended two meetings

in total, 11 (21.6%) attended three meetings in total, 17 (33.3%) attended 4 meetings in total; and 15 (29.4%) attended all five meetings. At least one staff member from a school attended a median of 4 of the meetings.

Our pre-specified definition of compliance at the pupil level, for use with the CACE analysis, was that both of the following criteria had to be met:

- there was representation from the pupil's school at all five training events; **and**
- their teacher indicated, on the post-intervention teacher survey, that for pupils in that class, on average, two reflections per pupil per week were completed.

These criteria were only fulfilled by 7 (13.7%) of the 51 intervention schools, corresponding to 218 (10.2%) of the 2,139 (2,133) intervention pupils included in the primary analysis for KS2 mathematics (reading). Formal CACE analysis was therefore not conducted as this level of compliance was thought too low to allow for a meaningful analysis.

Missing data analysis

Among those with non-missing KS1 mathematics score (n=4,221), a lower KS1 mathematics score was associated with being more likely to have missing KS2 mathematics data odds ratio (OR) for increase in 1 point in KS1 mathematics 0.74, 95% CI 0.69 to 0.78, p <0.001). FSM pupils were over twice as likely as non-FSM pupils to have missing outcome data (OR 2.16, 95% CI 1.28 to 3.65, p = 0.004), control pupils were nearly twice as likely as intervention pupils to have missing data (OR 1.83, 95% CI 1.05 to 3.20, p =0.03) but this was largely driven by the fact that KS2 mathematics outcome was not requested for three control schools as they withdrew from the evaluation. Pupils from schools in the South of England were over twice as likely as those in the North of England to have missing data (OR 2.36, 95% CI 1.19 to 4.69, p = 0.01).

Among those with non-missing KS1 reading score (n=4,220) a lower KS1 reading score was associated with being more likely to have missing KS2 reading data (OR for increase in 1 point in KS1 reading 0.74, 95% CI 0.70 to 0.78, p <0.001). FSM pupils were nearly twice as likely as non-FSM pupils to have missing outcome data (OR 1.80, 95% CI 1.05 to 3.07, p = 0.03).

Following multiple imputation, the adjusted mean difference in KS2 mathematics score was -1.13 (97.5% CI -4.08 to 1.83). The estimated Hedges' g effect size was -0.05 (97.5% CI -0.17 to 0.07), which relates to 1 month's less progress in the intervention group. No statistically significant difference was observed (p = 0.39). The total variance used to calculate the effect size was 601.45 the sum of 556.68 (random variation between pupils, within-cluster variance) and 44.77 (heterogeneity between schools, between-cluster variance).

Following multiple imputation, the adjusted mean difference in KS2 reading score was 0.10 (97.5% CI -0.96 to 1.16). The estimated Hedges' g effect size was 0.01 (97.5% CI -0.09 to 0.11), which relates to 0 months' additional progress in the intervention group. No statistically significant difference was observed (p = 0.84). The total variance used to calculate the effect size was 107.39; the sum of 99.76 (random variation between pupils, within-cluster variance) and 7.63 (heterogeneity between schools, between-cluster variance).

Subgroup analyses

<u>FSM in KS2 mathematics</u>

There was a statistically significant interaction between FSM status and allocation (interaction term 2.75, 95% CI 0.13 to 5.37, p = 0.04).

The primary analysis for KS2 mathematics was repeated in the subset of pupils with FSM status; the model included 1,330 pupils (702 intervention, 628 control), from 94 schools. The adjusted mean difference in KS2 mathematics score between the intervention and control groups among the FSM subgroup was 1.30 (95% CI -2.57 to 5.16) (Appendix Table 2). The estimated Hedges' g effect size was 0.05 (95% CI -0.10 to 0.20), which relates to approximately 1 month's additional progress in the intervention group (Table 13). No statistically significant difference was observed (p = 0.51). The total variance used to calculate the effect size was 669.33; the sum of 630.55 (random variation between pupils, within-cluster variance) and 38.78 (heterogeneity between schools, between-cluster variance). The ICC for KS2 mathematics score among the FSM subgroup associated with school was 0.15 (95% CI 0.11 to 0.20).

FSM in KS2 reading

There was no evidence of an interaction between FSM status and allocation (interaction term 0.42, 95% CI -0.66 to 1.50, p = 0.45).

The primary analysis for KS2 reading was repeated in the subset of pupils with FSM status; the model included 1,332 pupils (700 intervention, 632 control), from 94 schools. The adjusted mean difference in KS2 reading score between the intervention and control groups among the FSM subgroup was 0.47 (95% CI -0.96 to 1.90) (Appendix D Table 2). The estimated Hedges' g effect size was 0.04 (95% CI -0.09 to 0.17), which relates to 0 months' additional progress in the intervention group (Table 13). No statistically significant difference was observed (p = 0.52). The total variance used to calculate the effect size was 118.21; the sum of 111.19 (random variation between pupils, within-cluster variance) and 7.02 (heterogeneity between schools, between-cluster variance). The ICC for KS2 reading score associated with school was 0.09 (95% CI 0.06 to 0.14).

Table 11: Primary, secondary, and subgroup analyses

Outcome	Unadjusted means			Effect size			
	Intervention group		Control group				
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges' g (97.5%/95% CI)	p-value
KS2 Mathematics	2,217 (64)	74.8 (73.8, 75.9)	2,081 (164)	76.5 (75.4, 77.5)	4,148 (2139; 2009)	-0.05 (97.5% CI -0.18, 0.09)	0.43
KS2 Reading	2,209 (72)	32.4 (31.9, 32.8)	2,088 (157)	32.5 (32.1, 32.9)	4,149 (2133; 2016)	0.01 (97.5% CI -0.10, 0.11)	0.91
KS2 GPS	2,210 (71)	45.3 (44.7, 45.9)	2,087 (158)	46.5 (45.9, 47.1)	4,149 (2134; 2015)	-0.07 (95% CI -0.18, 0.04)	0.21
JrMAI	1,942 (339)	27.1 (27.0, 27.3)	1,758 (487)	27.2 (27.1, 27.4)	3,543 (1875; 1668)	-0.04 (95% CI -0.16, 0.07)	0.46
KS1 Mathematics	1,909 (412)	38.6 (37.9, 39.3)	1,841 (428)	39.1 (38.4, 39.7)	3,713 (1887; 1826)	-0.01 (95% CI -0.13, 0.11)	0.86
KS1 English	1,896 (425)	26.6 (26.1, 27.1)	1,751 (518)	26.8 (26.3, 27.3)	3,713 (1887; 1826)	0.03 (95% CI -0.08, 0.14)	0.57
KS2 Mathematics FSM subgroup	717 (35)	68.9 (67.0, 70.8)	642 (69)	67.3 (65.2, 69.3)	1,330 (702; 628)	0.05 (95% CI -0.10, 0.20)	0.51
KS2 Reading FSM subgroup	714 (38)	29.9 (29.1, 30.8)	646 (65)	29.4 (28.5, 30.2)	1,332 (700; 632)	0.04 (95% CI -0.09, 0.17)	0.52

Implementation and process evaluation

The IPE presented below is cross-sectional and longitudinal. Data was collected across the evaluation period from before randomisation (the baseline survey) up to and including June 2019 (at the end of the evaluation period). The IPE sought to:

- explore usual practice in the participating schools (both control and intervention) to evaluate what this means for the interpretation of the impact evaluation results—considering if this changed during the evaluation period and/or if there were any concurrent interventions taking place in control schools;
- explore the model of training applied where specific school staff (headteachers and lead practitioners) cascaded ReflectED training they received to the remainder of staff in their schools;
- assess the fidelity of the intervention (i.e. the extent to which it was delivered as intended in the intervention schools) and consider any barriers to implementation; and
- consider compliance data in relation to training (attendance at launch training and regional hub meetings) in addition to the compliance data that has been presented in the CACE analysis.

A full overview of the planned data collection for the IPE in relation to the research questions was presented in Table 3. What follows presents a summary of the data collected and analysed, by method; within this any changes from protocol or missing data are highlighted.

Summary of IPE data collected by method

The pre- and post-intervention surveys were conducted online. To minimise burden on schools, visits were not conducted in Spring Term 2018 to allow the schools time to establish their practice with ReflectED and where possible school visits incorporated all three of: structured observations of ReflectED; pupil focus groups; and teacher interviews. A total of 16 intervention schools were visited out of a planned 17 (one school cancelled and were unable to reschedule). The focus in the first year of the evaluation (Summer Term 2018) was on Year 1 and Year 5 pupils and on Year 2 and Year 6 in subsequent visits—this is reflected in the pupil focus groups, structured observations, and teacher interviews. Schools were selected using a range of different factors including school size, geographical location (e.g. urban or rural), and there were some re-visits planned in the second year of the evaluation as well as some visits to new schools.

Pre- and post-intervention surveys (headteachers and classroom teachers)

The baseline survey was administered between October 2017 and February 2018 and the post-intervention survey in July 2019. Table 14 provides an overview of the number of responses by allocation and teacher role at each time point.

Table 12: Survey data (pre and post) total number of responses

	Control schools		Intervention schools	
Respondent role*	Baseline survey (n)	Post-intervention survey (n)	Baseline survey (n)	Post-intervention survey (n)
Headteacher	45	39	46	46
Headteacher and nominated lead practitioner	5	0	4	1
Headteacher, nominated lead practitioner and KS2 teacher	0	0	3	0
Headteacher and KS1 teacher	1	1	0	0

Headteacher and KS2 teacher	3	3	3	0
Headteacher, KS1 and KS2 teacher	2	0	0	0
KS1 teacher	49	35	47	44
KS1 teacher and nominated lead practitioner	10	0	8	6
KS2 teacher	46	29	44	38
KS2 teacher and nominated lead practitioner	20	0	13	10
KS1 and KS2 teacher	0	0	0	2
KS1 and KS2 teacher and nominated lead practitioner	2	0	0	9
Nominated lead practitioner for ReflectED	11	0	12	31
Total (n)	194	107	180	179

*All participating schools were asked to nominate a lead practitioner for ReflectED prior to randomisation and as part of the MOU. Consequently, the control schools had a nominated lead practitioner for ReflectED at this point (baseline survey), but not when completing the post-intervention survey.

Interview data

The total number of teacher interviews conducted, transcribed, and analysed was 20, with 34 respondents in total:

- Summer Term 2018: six classroom teacher interviews in total, comprising five Year 5 teachers and five Year 1 teachers. Four of the interviews were joint (i.e. the Year 5 and Year 1 teacher together) and two were single teacher interviews (1 x Year 5 teacher and 1 x Year 1 teacher).
- Autumn Term 2018: seven classroom teacher interviews in total, comprising six Year 6 teachers and four Year 2 teachers. Three of the interviews were joint and four were single teacher interviews (x 3 Year 6 teacher interviews, and x 1 Year 2 teacher).
- Spring and Summer Terms 2019: eight teacher interviews in total, comprising six Year 6 teachers and seven Year 2 teachers. Four of the interviews were joint and four were single teacher interviews (three Year 6 teachers and one Year 2 teacher). In this group of interviews, four of those transcribed included a lead practitioner for ReflectED.

The trial protocol (Gascoine, *et al.,* 2018b) stated that interviews would be conducted with headteachers at the end of the evaluation period. It was not possible to arrange school visits that incorporated observations, pupil focus groups, teacher interviews, and headteacher interviews. Informal discussions with headteachers in the participating schools in the process of arranging visits informed the decision to only include headteachers in the post-intervention survey as the headteachers that we spoke to perceived that the most useful conversations we could have would be with classroom teachers rather than themselves. We considered the burden of IPE data collection on schools in this process, and it was necessary to make a decision about headteacher interviews that would maximise the data we were able to schedule collection of in selected schools (i.e. observations, pupil focus groups, and teacher interviews). Four of the eight teacher

interviews conducted between Spring Term 2019 and Summer Term 2019 included a lead practitioner. Lead practitioners had oversight of ReflectED in the whole school.

One semi-structured interview was conducted with the ReflectED developers, with two of the development team at the end of the evaluation period in October 2019.

Pupil focus groups

A total of 19 pupil focus groups were conducted. Each focus group was made up of 4–6 school selected students. Table 13 below shows the total number of focus groups conducted, with which year group(s) and when they happened.

Table 13: Pupil focus groups—data gathered

Term(s)	Year 1	Year 5	Year 2	Year 6	Total
Summer Term 2018 (May 2018 – July 2018)	2	4	N/A	N/A	6
Autumn Term 2018 (November 2018 – December 2018)[a]	N/A	N/A	2	5	7
Spring Term 2019	N/A	N/A	2	2	4
Summer Term 2019	N/A	N/A	2[b]	0[c]	2
Total	2	4	6	7	19

[a] One school visit in this term was cancelled by the school due to staff absence.
[b] A Year 2 focus group was conducted in one of these schools that was not planned for; the Year 6 teacher was absent on the day of the visit.
[c] One school cancelled a visit, there was not enough time left in the school year to arrange an alternative date.

Structured observations

In total, 16 structured observations of ReflectED were conducted in 14 different schools. In the Summer Term 2018 there were six observations (four Year 5 and two Year 1 lessons), in the Autumn Term 2018 there were seven observations (five Year 6 and two Year 2), in the Spring Term 2019 there were three observations (one Year 6 and two Year 2). A total of three planned observations in the Summer Term 2019 were cancelled by the participating schools due to unforeseen staffing issues.

There was a high level of agreement between observers in describing the delivery and engagement with the ReflectED lessons that were observed. The observation framework, which was agreed with the developers, was split into three sections: i) teacher behaviours; ii) child behaviours; and iii) whether there was evidence of performance tag colours displayed in the classroom, and whether Seesaw or reflection sheets were used to facilitate reflection at the time of the observation.

To explore any changes in the implementation of ReflectED (via structured observations of lessons) over the evaluation period, we had planned that four of the schools visited in the Summer Term of academic year 2017/2018 would be re-visited towards the end of the evaluation period, three of these re-visits were achieved with one cancelled by the school due to staffing issues.

Pupil reflections

Research question 13 asked ‘What evidence is there in pupils’ reflections of metacognition and any change in this over the course of the intervention?’ The planned approach to answer research question 13 involved an analysis of pupil reflections, completed as part of the ReflectED intervention on Seesaw. It was anticipated that the reflections would be

analysed using the Moseley, *et al.* (2005) framework, as had been done in previous research where pupils had ReflectED on their own learning using a visual research tool—Pupil Views Templates (Wall, 2008; Wall, *et al.,* 2007). However, it became apparent as we gathered evidence of student reflections that it was not going to be appropriate to analyse them in this way—the frequency of reflections collected in Seesaw was not as high as was expected in line with the logic model and intended delivery of ReflectED. The IPE data about barriers (research question 8b) explores this further and presents evidence from teacher interview data and the post-intervention survey to support this claim. Many of the reflections had been completed on paper rather than in Seesaw as we had originally anticipated, therefore it was not possible to analyse pupil reflections for evidence of metacognition in the originally planned way. Evidence of change in relation to evidence of metacognition in pupil reflections was instead derived from teacher-reported data from the post-intervention survey.

Having summarised the aims of the IPE and highlighted any changes from the protocol, what follows discusses the findings of the IPE in relation to each of the research questions posed. For clarity, the relevant research questions are listed under the section in which they are discussed. Presented at the beginning of each subsection is a list of the data sources used and a bullet point summary of the main findings in relation to the specified research question(s).

Usual practice

This section explores the IPE results in relation to usual practice for metacognition across all participating schools prior to randomisation (baseline) and within and at the end of the evaluation period (to understand if there were any concurrent interventions being used and if usual practice had changed in control schools). Data was gathered pertaining to usual practice in relation to metacognition, and for usual practice in relation to mathematics and reading (and literacy) given that these were the primary outcomes in the impact evaluation.

Data sources:

- Pre-randomisation baseline survey (all schools, control, and intervention)
- Post-intervention survey (all schools, control, and intervention)
- Teacher interviews (intervention schools)

Research question(s):

- 10: ‘How does ReflectED differ from existing practice in primary schools that focuses on facilitating pupils’ metacognitive skills?

Findings summary:

- There was similarity between existing practice in relation to metacognition at baseline across control and intervention schools, with control schools appearing slightly more ‘invested’ in metacognition and associated concepts at baseline.
- Usual practice in relation to metacognition at baseline included concepts like Growth mindset, this is a concept that is part of the training for the ReflectED intervention.
- Although metacognition was perceived, at baseline, as an important concept in the survey data analysed, there was most often not one specific approach to metacognition employed in the participating schools (rather a mixture of approaches). Linked to this, thematic analysis showed that there was often uncertainty about how metacognition was defined or what it looked like in relation to the practice of the survey respondents.
- Post-intervention there was some evidence in the data analysed of control school activity that had similarity, of varying levels, with aspects of the ReflectED intervention (e.g., specific time to reflect) and that teachers in control schools had engaged with metacognition related training. This was not unexpected, and reflective of usual practice and previous engagement with training (including metacognition related training) that was described at baseline.
- In terms of usual practice (at baseline) relating to the primary outcomes of mathematics and literacy approximately half of all participating schools reported using a whole school, KS1, or KS2 mathematics of numeracy intervention (49%) and slightly more reported using a whole school, KS1, or KS2 literacy intervention. These figures were lower for all groups in the post-intervention survey.

How important is metacognition?

At baseline, headteachers and classroom teachers were asked in the pre-randomisation survey how far they agreed with a series of statements regarding metacognition. The baseline and end of evaluation survey data presented in Figure 5 suggests that across both the control and intervention schools (prior to randomisation), metacognition was perceived as being an important concept, albeit slightly more so in schools subsequently allocated to the control condition than those allocated to receive the intervention. It is interesting to note that the perceived importance of metacognition for control schools had declined (79% to 57% as per Figure 5 below) and for intervention schools it had increased (70% to 92% as per Figure 5 below) by the end of the evaluation period.

Figure 5: Perceived importance of metacognition

Control schools

Baseline: 79% of control school headteachers (n=56) who completed the baseline survey reported that they 'strongly agree[d]' or 'somewhat agree[d]' that metacognition was an important concept in their school; and 82% of control school classroom teachers (n=142) 'strongly agree[d]' or 'somewhat agree[d]' that metacognition was an important concept in their classroom and teaching

End of evaluation: 57% of control school headteachers (n=42) who completed the post-intervention survey reported that they 'strongly agree[d]' or 'somewhat agree[d]' that metacognition was an important concept in their school; and 68% of classroom teachers (n=68) who completed the post-intervention survey 'strongly agree[d]' or 'somewhat agree[d]' that metacognition was an important concept in their classroom and teaching

Intervention schools

Baseline: 70% of intervention school head teachers (n=56) who completed the baseline survey reported that they 'strongly agree[d]' or 'somewhat agree[d]' that metacognition was an important concept in their school; and 75% of intervention school classroom teachers (n=131) 'strongly agree[d]' or 'somewhat agree[d]' that metacognition was an important concept in their classroom and teaching

End of evaluation: 92% of intervention school headteachers (n=48) who completed the post-intervention survey reported that they 'strongly agree[d]' or 'somewhat agree[d]' that metacognition was an important concept in their school; and 94% of classroom teachers (n=98) who completed the post-intervention survey 'strongly agree[d]' or 'somewhat agree[d]' that metacognition was an important concept in their classroom and teaching

Is specific time set aside for metacognition?

At baseline, headteachers and classroom teachers across all schools were asked whether specific time was currently set aside for metacognition in the current approach to metacognition of their school or classroom. A total of 11 headteachers (of 112 respondents) and 38 classroom teachers (of 271 respondents) reported that specific time was set aside (timetabled) in their school or classroom for metacognition: 7 headteachers in control schools; 4 in intervention schools; and 18 classroom teachers in control schools and 20 in intervention schools.

At the end of the evaluation this question was asked of control school classroom teacher respondents (n=67) again, responses showed that there had been a change—8% (n=12; compared to 13% at baseline) of classroom teachers reported that specific time was set aside (timetabled) in their school or classroom for metacognition.

Training related to metacognition

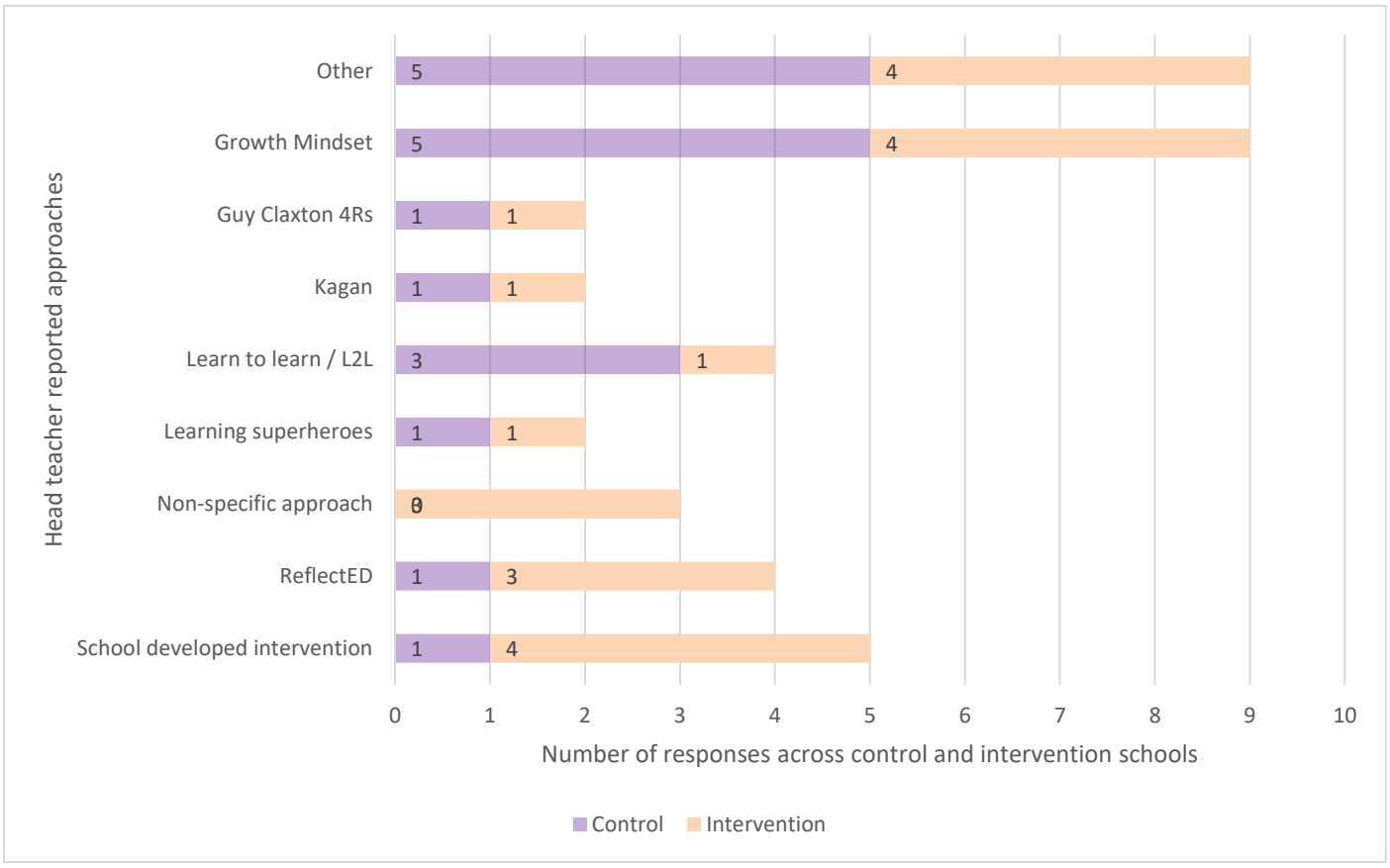
At baseline, headteachers across all participating schools were asked if they or their staff (to the best of their knowledge) had received any training specific to metacognition or related concepts, and also to provide details about the training they had received if they answered 'yes'. In total, 32% of the headteachers (36 headteachers out of 113 respondents) surveyed at baseline reported that they or their staff (to the best of their knowledge) had received training specifically relating to metacognition. In total, 30% of the classroom teachers surveyed (84 out of 278 respondents) reported that they had received training that was specifically related to metacognition, 36% subsequently in the control condition (51 out of 143) and 24% allocated to receive the intervention (32 out of 131).

At the end of the evaluation period, Year 2 and Year 6 teachers in the control group were asked about any training related to metacognition and when they received this training. In total, 49% of respondents (33 out of 68 respondents) reported that they had received training, additionally 48% of those who reported having done so indicated that this was in the last five school terms (i.e. during the intervention period) (16 out of 33 respondents). The number of control school teachers reporting having received training related to metacognition was higher at follow-up than baseline (49% compared to 36% at baseline; 33 and 51 respondents, respectively), suggesting that there was some activity in control schools related to metacognition.

Whole-school approaches to metacognition

At baseline, headteachers were asked if their schools were currently implementing any whole-school approaches to metacognition, 34% (n=19) of the control school headteacher respondents (n=56) reported that they were. Those that answered 'yes', were asked to provide brief details of the approach in a free-text response, that were coded inductively using NVivo and the results of which are presented below. Some headteachers gave information about more than one whole-school approach to metacognition, if this was the case all the approaches listed were coded. Figure 6 illustrates the 31 free-text responses coded, those approaches that were coded more than once (those that were not, are amalgamated in the 'other' category).

Figure 6: Headteacher reported whole-school approaches to metacognition (baseline)



As Figure 6 shows, the most common headteacher response was 'growth mindset' (29% of the responses in total) being described as the whole-school approach to metacognition that their school implemented. Growth mindset does cross over with the training delivered in relation to the ReflectED intervention, as was observed at ReflectED training sessions. As Figure 6 shows, a small number of headteachers (five in total: four intervention and one control school) described 'school developed intervention[s]' as the approach(s) to metacognition that their school used. Examples of headteacher descriptions of school developed intervention included:

> *… this is an in school approach where children reflect upon their learning focusing on areas such as what knowledge did I know that helped me, what strategy did I use that helped me, how could I improve on my learning experience. (Headteacher, intervention school)*

and… *we have developed our own 'learning powers' which we teach from year 1 upwards – however this is an early part of the process for us. (Headteacher, intervention school)*

Headteachers were asked to specify who delivered the whole-school approach to metacognition that they gave details for and who received it. Around 65% of the respondents (20) did not report a specific person (e.g. headteacher, class teacher) or group of people (e.g. classroom teachers) as responsible for delivering the whole-school approach to metacognition that they detailed. Of those headteachers that did report a specific person(s) as responsible for delivery: two were higher level teaching assistants (HLTAs) (both in intervention schools); one teaching assistant (control school); eight teachers (four each in intervention and control schools); and two 'whole school' (i.e. all staff).

A total of four headteachers reported that they were already implementing aspects of ReflectED in their school (three intervention and one control), but that they had not received individualised training or support in relation to it. Given that prior to the evaluation ReflectED resources were freely available online, we did not state that schools who had previously engaged with ReflectED resources would be excluded from the evaluation. However, headteachers and classroom teachers were specifically asked about any previous engagement with ReflectED resources in the baseline survey. Around 7% of headteachers (eight respondents) stated that they or their school had engaged with ReflectED resources, equally balanced across both intervention and control conditions, as did 6% of teachers (8% control and 3% of intervention). This tended to primarily involve attending training, mainly within school or locally (reported by 5 out of 8 headteachers and 5 out of 16 teachers) and the use of online resources (reported by 3 out of 8 headteachers and 10 out of 16 teachers). Online resources were particularly reported by control schools although the numbers are so small as to make comparison meaningless. In addition, one of the respondents from the intervention group reported that ReflectED resources had been trialled previously in Year 1 after the headteacher visited Rosendale Primary School.

During the semi-structured interviews (within the evaluation period), teachers were asked to describe their school's approach to metacognition prior to the implementation of ReflectED. There was a range in teacher's responses. Most teachers interviewed reported that they did not have specific metacognition lessons prior to ReflectED, and in some cases, metacognition was also not part of their teaching practice at all. This aligns with the survey data (low numbers reporting specific time set aside) and the thematic analysis of the survey data relating to 'what does metacognition look like' presented below.

Some teachers discussed that the extent of metacognition teaching was the use of buzzwords, such as 'resilience' and 'perseverance'. Other teachers reported metacognition teaching practices and strategies that bore similarity to those taught within ReflectED such as asking children to self-reflect after lessons, Kagan's tables, mixed ability pairs, and growth/fixed mindset. One school noted that they had already been using Seesaw prior to their participation in the trial. However, a key theme to emerge from the interviews was that all schools lacked an effective and practical method of teaching metacognition.

> *As a school, we tried historically, we had what we called our values, really, which were things like resilience, and perseverance. So we had core things that we tried to teach through but we didn't have a structured way in teaching it, so when we heard about … we saw one of the teachers who was talking about what they'd done, and we were really interested because we thought it was a structured way to bring it into school. (Classroom teacher, intervention school)*

All schools recognised the importance of metacognition for children's learning and foresaw value in acquiring a structured method of teaching it and the potential benefit it could have on their pupils in terms of raising their awareness of an involvement in the process of their own learning. For example:

> *I know that we're taking part in it because there's evidence that children who are more involved in their own learning process do better, because they can identify what it is they're good at, what it is they're not so good at and then they can work on those areas. So that's why we've introduced it. (Classroom teacher, intervention school)*

These reasons, coupled with a lack of current structured teaching for metacognition, were often cited as the driving force behind a school's decision to participate in the trial.

The post-intervention survey also asked headteachers in the control schools (n=43) if any metacognition focused interventions were currently being implemented. Around 10% (n=4) of headteachers who responded to this question

reported that metacognition focused interventions were currently being implemented in their school, with 90% (n=38) indicating that there were not. Of those who were implementing a metacognition intervention, one control school explicitly mentioned using ReflectED. In contrast, 71% (n=30) of headteachers in the control condition reported that they intended to implement ReflectED in the following academic year and the remaining 29% who answered this question indicated that they may do so (n=12).

Metacognition focused practice

Headteachers and classroom teachers were also asked, at baseline, how far they agreed with four statements that were focused on aspects of practice that link to metacognition and aspects of the ReflectED intervention:

- 'Pupils are confident in talking about their own learning';
- 'Pupils use appropriate vocabulary to talk about their learning';
- 'Pupils are confident to work independently'; and
- 'Pupils in my school have conversations about learning with each other'.

The responses are presented below in

Figure 7 (headteachers) and Figure 8 (classroom teachers) for baseline and showed that both headteachers and classroom teachers were largely in agreement about the importance of the aspects of pedagogy described in the statements.

Figure 7: Headteacher agreement at baseline (pre-intervention survey) with statements about talking about learning, appropriate vocabulary, and independent working (n=112 responses)

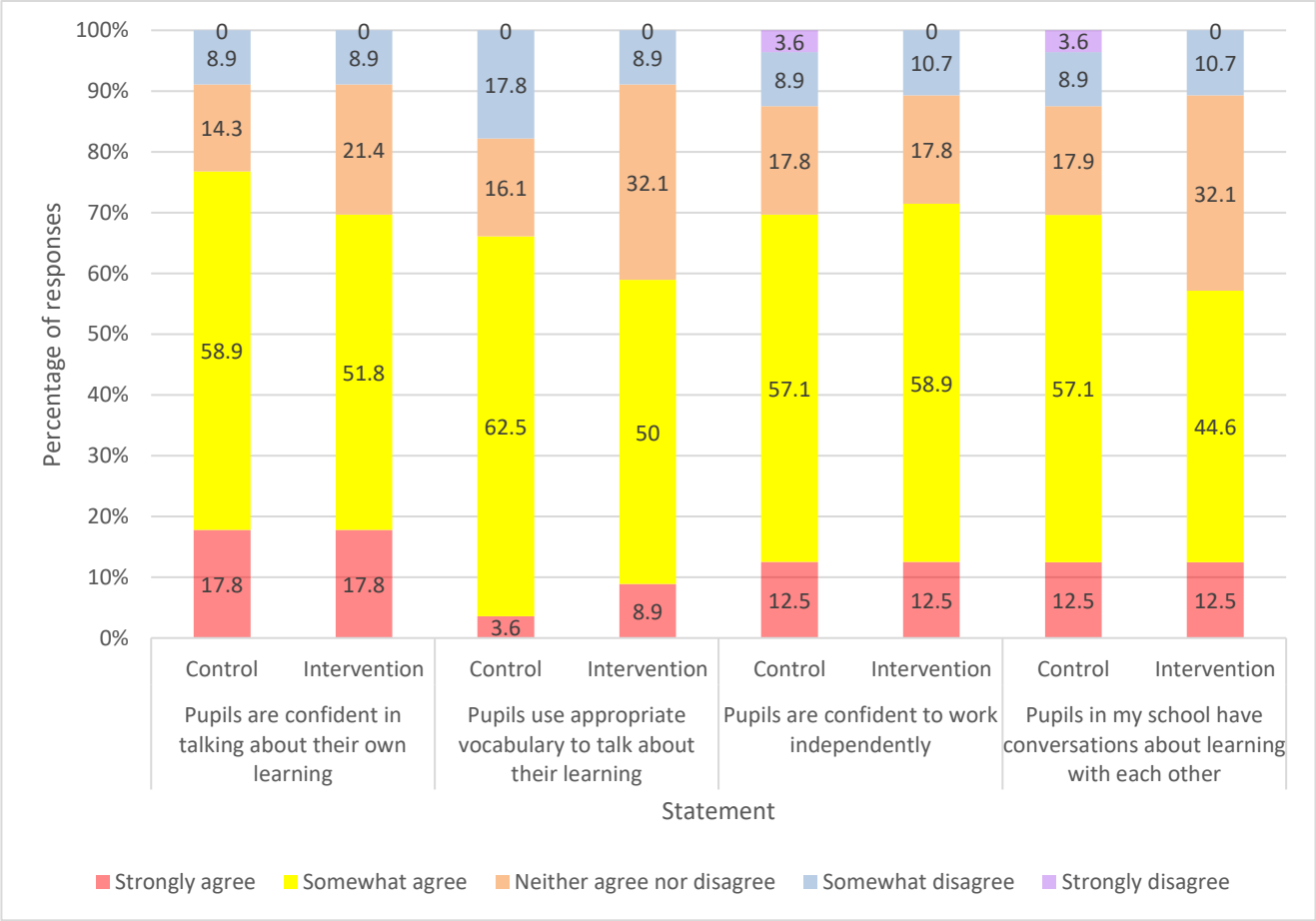
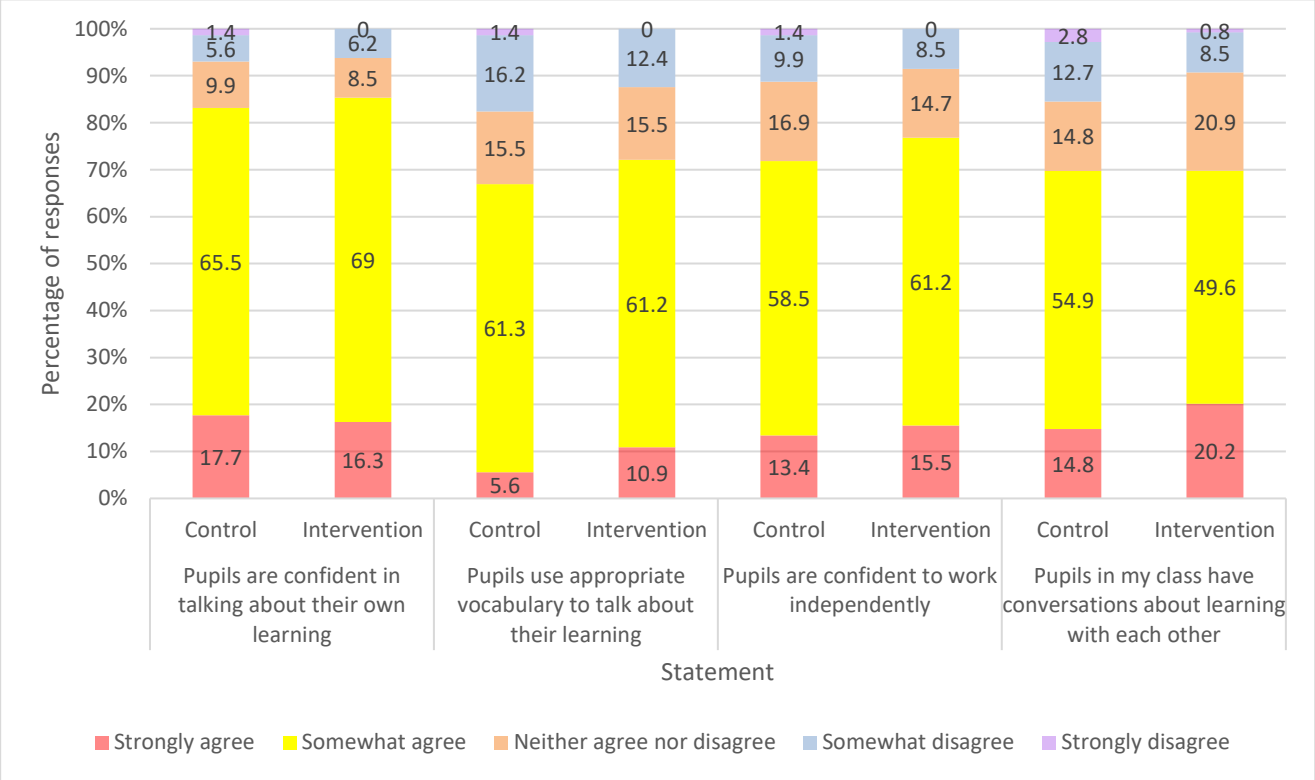


Figure 8: Classroom teacher agreement at baseline (pre-intervention survey) with statements about talking about learning, appropriate vocabulary, and independent working (n=217 responses)



The post-intervention control school survey also asked classroom teachers to respond again to the statements in Figure 8 above. There were increases across all of the statements:

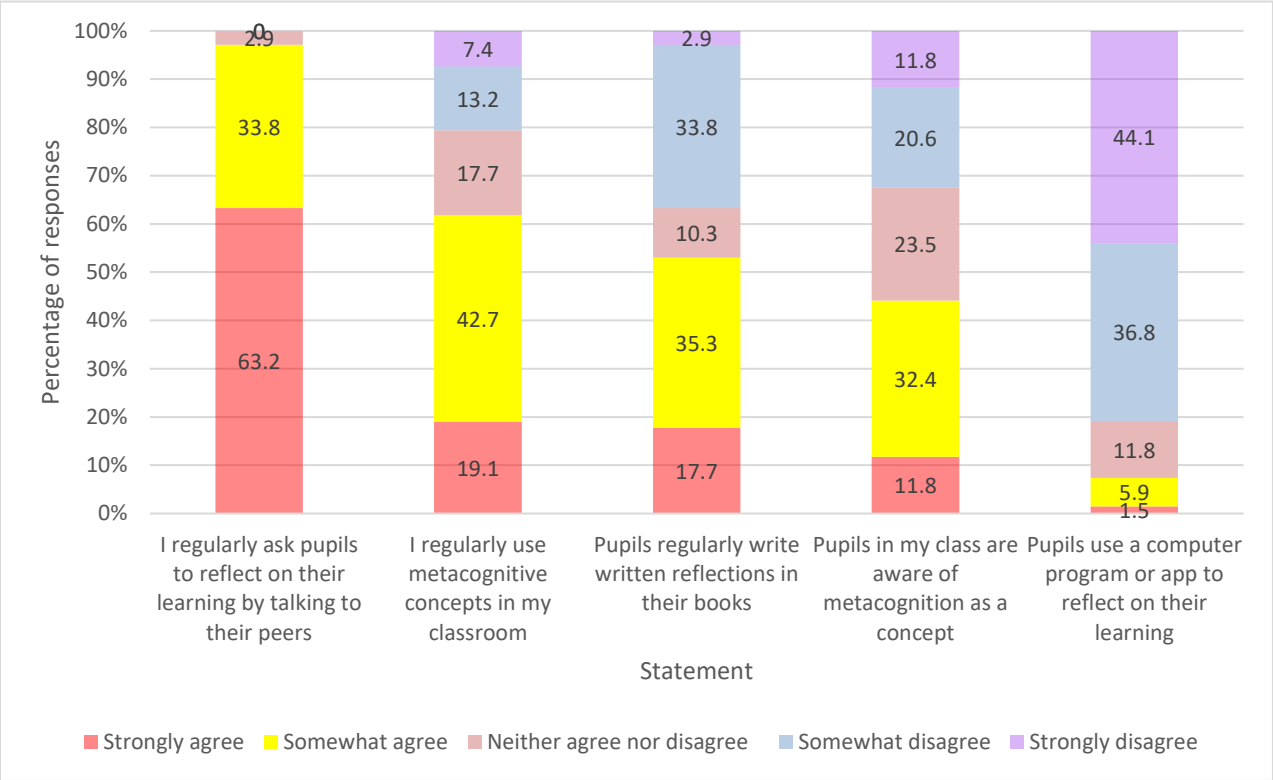
- 'Pupils are confident to work independently': strongly agree (pre- 13.4%, post- 33.8%; 19 out of 142 respondents and 23 out of 68 respondents, respectively);
- 'Pupils use appropriate vocabulary to talk about their learning': strongly agree (pre- 5.6%, post- 25%; 8 out of 142 respondents and 17 out of 68 respondents respectively);
- 'Pupils are confident in talking about their own learning': strongly agree (pre- 17.6%, post- 42.6%; 25 out of 142 respondents, and 29 out of 68 respondents respectively); and
- 'Pupils in my class have conversations about learning with each other': strongly agree (pre- 14.8%, post- 38.2%; 21 out of 142 respondents and 26 out of 68 respondents respectively).

The increase in 'strongly agree' responses above suggests that there were increases in aspects of practice that ReflectED addresses as part of the usual practice in control schools. What this comprised was explored in more detail in the control school post-intervention teacher survey as per Figure 9 below. The post-intervention survey for control schools asked classroom teachers (Year 2 and Year 6 teachers) about aspects of metacognitive pedagogy that were specifically linked to aspects of ReflectED to evaluate the extent to which there was similarity (with ReflectED) in the usual practice of control schools during the intervention period. This data is presented below in Figure 9. Of the 68 respondents, most control teachers at post-intervention:

- 'Agreed' (strongly or somewhat) that they regularly asked pupils to reflect on their learning by talking to their peers (97% or 66 teachers);
- 'Agreed' (strongly or somewhat) that they regularly used metacognitive concepts in their classroom (62% or 42 respondents); and
- 'Agreed' (strongly or somewhat) that their pupils regularly wrote reflections in their books (53% or 36 teachers).

Just under half (44% or 30 teachers), however, agreed ('strongly or somewhat') that the pupils in their class are aware of metacognition as a concept with a further 32% disagreeing (either 'somewhat or strongly') with this statement. Only 7% agreed ('strongly or somewhat'; five respondents) that their pupils used a computer program or an app to reflect on their learning. Of these 7% (four respondents) reported using Seesaw, which is also used as part of ReflectED.

Figure 9: Control school classroom teacher (n=68) reported aspects of metacognitive practice (post-intervention)



To summarise, activity in control schools during the intervention period does have overlap with aspects of the ReflectED intervention including regularly asking pupils to reflect on their learning and to complete written reflections.

What does metacognition look like?

In terms of describing what existing metacognitive practice looked like at the time of gathering the baseline survey data (pre-randomisation) for all the participating schools both headteachers and classroom teachers were asked to respond to the following statements that requested an open-ended response by way of description:

- Baseline survey (headteachers): 'Please describe what metacognition looks like in your school' (i.e. specific practices in terms of teaching and pedagogy, language used to talk about learning, specific approaches, or interventions …) (n=110 responses analysed, 55 intervention, 55 control).

- Baseline survey (classroom teachers and lead practitioners): 'What does metacognition look like in your classroom?' (e.g. Do you follow specific guidance directed by the school, do you set aside time in your lessons or in specific lessons? Do you follow any specific practices or use specific language to talk about metacognition and learning?) (n=263 responses, 137 intervention, 126 control).

A reflexive thematic analysis approach (Braun and Clarke, 2021) was used to analyse this free-text data (that incorporated a total of 373 free-text responses from distinct respondents) so as to capture the depth and richness of it, in describing what metacognition looked like in evaluation schools and classrooms prior to involvement within this evaluation. The focus and intention of the thematic analysis was on the research question posed (research question 10)—undoubtedly the large number of free-text survey responses provided a rich source of data that in many cases went far beyond the question of what metacognition looked like currently (in the respondent's schools or classrooms). What is presented here focuses on the overarching themes and themes that are relevant to the research question posed. The two overarching themes were: i) *metacognition is new;* and ii) *metacognition is linked to what we already do*, their

characteristics and those of their underlying themes are presented below. The overarching themes and themes were common across the dataset (both control and intervention schools), they are not differentiated in the illustrative thematic analysis results and discussion that are presented below. The overarching themes and the themes that underlie them are presented in Table 16.

Table 14: What does metacognition look like? Thematic analysis of survey responses

Overarching themes	Themes
Metacognition is new Perceptions of metacognition where there was a sense of newness in the school and/or for the individual respondent. Sometimes newness was explicitly stated and other times it was more implicit and focused on the individual (e.g. an individual not having had any formal training)	I am / we are unsure Strong sense of being unsure about how to define and/or understand what metacognition is and what this looks like in practice
	I am / we are exploring A sense of exploration within schools and individual classrooms of what metacognition is and what it looks like in a setting. Sometimes there was a suggestion that metacognition was beginning to be embedded
Metacognition is linked to what we already do Respondent reported perceptions of metacognition were described in relation to existing practice and pedagogy	Talk and language about learning Description of metacognition that focused on talking and dialogue about learning, between pupils, and between pupils and teachers (often modelling for the latter)
	Reflecting on learning The role of pupil reflection in what metacognition looks like in schools and classrooms. Often this focused on reflecting on the process of learning
	Specified interventions or approaches More confident descriptions of metacognition that focused on specific intervention(s), these were a mixture of named (and recognisable interventions) and school developed approaches

The thematic analysis illustrated how respondents regularly expressed the notion that metacognition was a new and sometimes unfamiliar concept, some respondents commented that there was no metacognition in their school setting and/or classroom. Some respondents acknowledged their unfamiliarity with metacognition, or a difficulty in defining it was apparent in the response (e.g. there was a sense of tentativeness or questioning if they had the right answer in the response). This contrasted with others who seemed confident in their description of what metacognition looked like and there was not a sense that it was a new concept, indeed it was described in relation to existing practice in schools and classrooms. Consider the response of one headteacher from a school later allocated to the intervention:

> *Metacognition is a new concept for [name of school redacted]. Whilst staff have a good understanding of the importance of growth mindset in ensuring children feel confident to tackle challenges, they struggle to describe the necessary metacognitive skills needed in order to achieve their next steps in learning. The staff struggle to model to the children what the learning process looks like, although some have limited knowledge of 'visible learning' strategies such as being in the 'learning pit' and how to move forward with difficult concepts.*

While metacognition is described explicitly as new in the above extract, understanding of other specified interventions or approaches (i.e. Growth Mindset, the idea of visible learning) is clear and asserted as something where there is already a developed understanding. This understanding and confidence in relation to Growth Mindset is positioned alongside a ‘struggle’ in relation to metacognitive skills.

Some described that there was no metacognition in their school or classroom in quite definite terms:

> *There is nothing currently in place. (Headteacher, intervention school)*

> *No evidence of this being taught or discussed in my school. (Headteacher, intervention school)*

However, overall, there is a sense in the data that difficulty in defining metacognition was not uncommon and that respondents were often unsure about whether or not what they described was metacognition or not:

> *I'm not sure if these are Metacognition things but I always talk to the children about what we are learning and why. (Teacher, intervention school)*

> *I would like some assistance with this. (Classroom teacher, control school)*

> *We are just starting on a journey to get metacognition into our teaching and learning in school. (Headteacher, intervention school)*

Sometimes respondents described that there was no metacognition in their classroom in quite definite terms and other times there was an air of uncertainty not dissimilar to that described above:

> *I'm not yet sure what it is so can't say whether I use it! (Classroom teacher, intervention school)*

> *I don't think that there is any evidence of metacognition in my classroom yet. (Classroom teacher, control school)*

The uncertainty around and/or lack of metacognition in some respondent's responses captured within the overarching theme of 'metacognition is new' is cognisant with the idea that schools perhaps signed up to participate in the evaluation of ReflectED because they were interested in and aware of the benefits of metacognition and looking for ways to explore and implement this within their settings. Some respondents, rather than describing an uncertainty, focused more on the notion of exploring—being at the beginning of engagement with metacognition, testing the water and in a process of considering what worked in their school and/or their classroom. One used the phrase *'we are just dipping our toes into work on metacognition' (headteacher, control school),* whereas many others communicated more of an enthusiasm for active exploration in relation to metacognition:

> *We have done quite a bit of work around effective questioning, and although other practice is present, (peer marking, peer feedback) it is not consistent or even prevalent throughout school. We want to embark on this key area of development this academic year. (Headteacher, intervention school)*

> *This is not specific guidance from the school, but as a school leader it's something I'd like to see in other classrooms and as part of our 'culture'. (Classroom teacher, intervention school)*

> *No specific interventions as yet (this is why we're interested in being part of the research) although we have had some training in growth mindset (2015–16) and staff encourage the idea of drafting and improving work using the concept of Austin's butterfly and the use of a 'purple polishing pen' to improve work. (Headteacher, control school)*

> *No specific guidance directed by the school or specific time in lessons to focus on metacognition. No specific practices used but we are trying to encourage the use of language linked to metacognition and learning. We are currently in the process of exploring/researching metacognition. (Classroom teacher, control school)*

It seems that there is an air of 'fuzziness' in respondents describing what metacognition looks like that is not dissimilar to the assertion that metacognition is a 'fuzzy' concept (Wellman, 1985). This is not unexpected given the plethora of different definitions of metacognition that are described in the research literature (as explored in the background section).

As well as a strong sense of metacognition being something new, for many of the participating schools, there was also a sense in the thematic analysis that many were able to make links between metacognition and what they already do. This aligns with the data from the survey where headteachers were asked to describe any existing whole-school approaches to metacognition that were already being implemented. Growth mindset was the most common in that data

and in this thematically analysed data, interventions that headteachers and/or classroom teachers specified as or in relation to metacognition (including Growth Mindset and others like Learning Power, Bloom's Taxonomy, Kagan, Philosophy for Children [P4C], Learning 2 Learn) were common. As well as commonality here, respondents often described explicitly how this was/these were existing practice in their setting prior to their involvement in the evaluation process:

> *In our school, we strongly believe in questioning the children and getting them to think a little deeper about what they are learning. In the classroom, this can be as simple as asking Bloom's Taxonomy style questions to develop learning. (Classroom teacher, intervention school)*

> *Pupils are used to working with Learning Partners in the school and classes use the Kagan model to support collaboration and discussion. Pupils are used to using Learning Partners to help each other to learn and we are currently looking at how we can teach pupils to use the Scaffold Framework to facilitate effective peer learning. (Headteacher, control school)*

Within the descriptions of existing practice that respondents offered as responses to the question of what metacognition looks like, as well as specified intervention or approaches, more generally responses relating to reflection and talk and language about learning were prominent. It is important to note the link, particularly in relation to reflection and language focused on learning, to the ReflectED intervention and logic model (Figure 1). These aspects of the ReflectED intervention were commonly described as what metacognition looked like (in schools and classrooms) across the participating schools prior to randomisation. For example:

> *While we don't teach metacognition in an isolated fashion, we do talk a lot about how we learn, how to be reflective, how to be resilient in learning and growth mindset. (Headteacher, intervention school)*

> *Children talk about what they learned and how, what made it easier or harder and add more to others ideas. They regularly offer 'what I could say or show to help you understand'. This happens in most lessons. I try to 'think out loud' while modelling. (Classroom teacher, control school)*

Looking at the latter example in more detail, the theme of reflecting on learning can be seen alongside talk and language about learning. Almost like these themes are inextricably linked, metacognition looks like something that includes reflection alongside the ability to be declarative about the process of learning in talking about it (or considering the steps whether this is verbalised or not).

The post-intervention survey also asked teachers in the control schools a more open-ended question about what metacognition looked like in their classrooms. In total, 65 free-text responses were analysed, the analysis looked specifically and in a deductive manner for aspects of metacognitive practice in the control schools that had crossover with the ReflectED intervention.

Of the 65 free-text responses, seven respondents recorded an answer that explicitly said that there was no metacognition in their classrooms or replied with text including 'N/A'. For example*: 'As a controlled school, we have not yet implemented any strategies into our classroom' (Classroom teacher, control school).* Reflection was mentioned in 40% of the responses (n=26), one of the key aspects of the ReflectED intervention is reflecting on learning and this as a metacognitive skill. However, we cannot claim that by virtue of this that control schools were implementing an aspect of ReflectED, this would not be accurate, and reflection has been considered widely in relation to pedagogy and within literature that is focused on metacognition (e.g. Desautel, 2009). Other aspects of the ReflectED intervention that appear to have been present in some of the control schools, as self-reported by teachers in the post-intervention survey, include using different colours as part of reflecting on learning (n=4) and growth mindset (n=3). Again, these could not be wholly described as classroom practice distinct to the ReflectED intervention. Language about learning and talking about learning, were also highlighted in these responses in the post-intervention survey for control schools as being common practice; a total of 16 responses considered that what metacognition looked like in their classroom incorporated talk and language about learning. For example:

> *We use a colour coded system of self-assessment which encourages the children to reflect on their learning. This is a whole-school approach, consistent in all classrooms. Sentence stems and key questions are used to encourage children to use appropriate vocabulary and structure when discussing their learning. Now the children are becoming more confident and the practice is embedded, less*

children require sentence stems and can reflect independently. Initially, and now occasionally, we spent lessons specifically learning how to think metacognitively and to develop metacognition skills and understanding. Now the children understand it, we are able to do 'little and often' five/ten minute chunks of reflection in almost every lesson. (Classroom teacher, control school)

We haven't introduced the meta-cognition concept to our children (as we were deemed to be a control school and so didn't want to disturb the study). However, we have continued to use talk partners, encouraged discussion and written feedback and have discussed with the children how best they learn to ensure we are presenting learning and feedback in the best way. (Classroom teacher, control school)

The two examples above are particularly salient examples and were selected because they incorporate different aspects of practice in the control schools where there is crossover, to varying extent, with aspects of the ReflectED intervention and seem to have been present in the control schools throughout. There was no evidence in the self-report survey that control schools introduced these 'metacognitive practices' specifically during the intervention period.

Additional usual practice data: literacy and numeracy interventions

Additional data pertaining to usual practice gathered focused on the primary and secondary outcomes of mathematics and reading and is summarised below. Although the ReflectED intervention is focused on metacognition and not mathematics or reading it was important to understand if there were any mathematics or reading focused interventions running parallel to the intervention or control conditions as these were the primary outcomes.

Baseline survey:

- Nearly half (49%; 56 out of 114) of all headteachers reported that their school was currently implementing a whole school, KS1, or KS2 mathematics or numeracy intervention(s) e.g. Times Table Rock Stars, Mathletics, Maths Mastery.
- 56% of headteachers (64 out of 114) reported that their school was currently implementing a whole school, KS1, or KS2 literacy intervention(s) e.g. Read, Write Inc., Time to Talk, Pirate Crew, Pearson's Bug Club.
- In addition, all headteachers who reported implementing whole school, KS1, or KS2 mathematics/numeracy interventions of literacy interventions also indicated that one to one or small group support was also offered to individual children believed to be struggling with aspects of their mathematics/numeracy or literacy.

Post-intervention survey:

- By the end of the intervention period the proportion of headteachers reporting that their school was currently implementing a whole school, KS1, or KS2 mathematics intervention had decreased, particularly for intervention schools (45% of headteachers in the control condition, and 31% in the intervention condition; 19 out of 42 and 15 out of 48, respectively).
- Similarly, the proportion of headteachers reporting that their school was currently implementing a whole school, KS1, or KS2 literacy intervention(s) decreased, although with less variability between the conditions (46% of headteachers in control schools compared to 50% in intervention schools; 17 out of 43 and 24 out of 48, respectively).
- Where mathematics/numeracy or literacy interventions were reported they remained similar in nature to those reported in the baseline survey.

Compliance

Compliance data presented here should be considered alongside the CACE analysis presented in the impact evaluation; it considers the amount of ReflectED that has been delivered. Linked to this section is the section subtitled 'Training', data pertaining to compliance with (including attendance at) the training and fidelity associated with the ReflectED training is presented in this subsection.

Data sources:

- Secondary data gathered from the developers (registers of attendance at training events and hub meetings)
- Post-intervention survey (intervention schools only)—the frequency at which teachers taught ReflectED lessons and the average numbers of reflections that they reported pupils completing.

Research question(s):

- 9: How much of ReflectED has been delivered in the intervention schools?
- 11: 'What is the reach of ReflectED across the intervention schools (e.g., … how many reflections have been completed? ...)

Findings summary:

Training (both launch and regional hubs) were well attended. In terms of delivering the five school terms of ReflectED for the intervention period, most schools (86%) reported (via the lead practitioner in the post-intervention survey) that they began delivering ReflectED in Spring Term 2018 as was expected. Referring to the logic model (Figure 1), we note the importance of reflecting regularly as part of the ReflectED intervention, with the recommendation being two reflections per pupil per week—compliance collected data (via the post-intervention survey) in relation to reflections shows that on average, as reported by classroom teachers, pupils were not completing the recommended two reflections per week.

In addition to the training compliance data that is explored in the 'Training' section, attendance can be summarised as follows. Attendance at the initial launch training was achieved by 54 (96.4%) of the 56 intervention settings.

- Attendees of the launch training were very positive about the usefulness of this training and the extent to which it equipped them to deliver ReflectED training in their own schools.
    - the post-intervention survey data (n=177 respondents) further supported this, with respondents reporting that the training was pitched correctly, prepared them to deliver ReflectED and that they were confident in understanding the principles of the ReflectED approach to metacognition. The post-intervention survey data refers to all training provided for the intervention, unless specified, some respondents attended launch training and others received the subsequent cascaded training within their school.
- Regional hub meetings were largely well attended.
    - headteachers largely reported that it was unproblematic for staff to be released up to attend regional hub meetings; and
    - regional hub attendees valued these meetings as supportive places to share practice and ask questions.

In the post-test survey the nominated lead teachers in the intervention schools were asked when their school started to use ReflectED. The majority (92% of 48 respondents) stated that their school started delivering the intervention in Spring Term 2018, after the initial training event: 38% (18 respondents) in January 2018; 44% (21 respondents) in February 2018; and 10% (5 respondents) in March 2018. Around 4% (two respondents) stated that they started using ReflectED later in the school year, 2% in May 2018, and 2% in July 2018. Surprisingly, 4% (two respondents) reported that their school started to use ReflectED in September 2018, this means that they delivered less of the intervention—three school terms, or one school year as opposed to the anticipated five school terms.[5]

Year 2 and Year 6 teachers in intervention schools were asked their agreement or otherwise about the ease of fitting ReflectED lessons into their weekly plans/schedule:

Nearly two-thirds (63%, 62 out of 98) of respondents agreed that it was easy to fit ReflectED lessons into their weekly plans/schedule (26% strongly agreed and 38% somewhat agreed; 25 and 37, respectively), 7% (7 respondents) neither agreed nor disagreed, 23% (23 respondents) somewhat disagreed, and 6% (6 respondents) strongly disagreed.

[5] Where there was disagreement or uncertainty on the start date, the researchers examined the data more closely to reach agreement on the month/year in which the school began delivering the programme (e.g. by cross-referencing with other survey responses).

Finally, Year 2 and Year 6 teachers were asked how many reflections were completed on average, per pupil per week. Just over half (54%; 52 out of 95) of respondents to this question indicated that on average pupils were completing one a week, 4% (4 respondents) indicated less than one a week (1% none, 3% 0.5 reflections; one and three respondents, respectively; where more than one number was given this was averaged out), 11% (10 respondents) indicated 1.5 reflections completed per pupil per week, 15% (14 respondents) 2 reflections per pupil per week, 3% (3 respondents) 2.5 reflections per pupil per week, 5% (5 respondents) 3 reflections per pupil per week, 1% (1 respondent) 4 reflections per pupil per week, 2% (2 respondents) 5 reflections per pupil per week, 1% (1 respondent) 6.5 reflections per pupil per week, 1% (1 respondent) 10 reflections per pupil per week, and 1% (1 respondent) 14 reflections per pupil per week.

Training

Data gathered in relation to training in the IPE relates to both compliance (as explained in the previous section) and fidelity (i.e. when considering how effective the cascaded model of training has been). With this in mind, this section ('Training') is situated between the compliance and fidelity sections of this report.

Data sources:

- Training registers kept by the development team—compliance;
- Post training survey for training attendees—compliance and fidelity;
- Post-intervention survey—compliance and fidelity;
- Training structured observations—fidelity;
- Teacher interviews—fidelity; and
- Developer interview—fidelity.

Research question(s):

8a. '([...]How effectively has the training provided to the headteachers and the lead practitioners cascaded to

the remaining teachers?')

11. 'What is the reach of ReflectED across the intervention schools? For example, What proportion of training

has been attended? …'

Findings summary:

Data presented in this section shows a high level of compliance to the intended model of training and cascade for the ReflectED intervention and would suggest that the training provided to headteachers and lead practitioners was effective in that they were then able cascade it to the remaining teachers in their schools who reported that it was well received. Considering the reach of ReflectED in relation to training attendance the data presented in this section illustrates a high level of attendance across the intervention schools.

The ReflectED training model and training attendance

The training model for ReflectED focused on the delivery of training from the ReflectED development team to headteachers and nominated lead practitioners from each intervention school. The headteachers and lead practitioners were then required to cascade the training that they had received to the remaining staff in their schools. Lead practitioners were also required to attend termly regional hubs that were delivered by the ReflectED development team. Training was observed by two members of the evaluation team (LG and LT); they attended one iteration of the initial 'launch' training event for intervention schools of which there were five iterations that took place at Rosendale Primary School between 11 January 2018 and 8 February 2018. LG and LT also attended five network hub meetings between them between May 2018 and March 2019.

LG and LT also attended a ReflectED training day prior to the evaluation period in June 2017 during the set up stage of the evaluation. Attending this training day gave them a working knowledge of the intervention in action at Rosendale Primary School and informed their subsequent development of the IPE materials.

Table 17 below presents the attendance data for training that was gathered by the developers for the launch training and the termly regional hubs. It should be noted that schools were expected, as per the MOU, to send the headteacher and nominated lead practitioner for ReflectED to the initial launch training. Two schools did not send any teachers to the

initial launch training. Headteachers from 8 of the remaining 54 intervention schools did not attend the initial launch training for ReflectED; 2 schools only sent 1 teacher (both lead practitioner, 1 of which was also the school's headteacher), 51 sent 2 members of staff to the training, including their nominated lead practitioner for ReflectED, and 1 sent 3 members of staff.

Two intervention schools (3.6%) did not attend any of the hub meetings, 2 (3.6%) attended 1 of the hub meetings, 5 (8.9%) attended 2 of the hub meetings, 12 (21.4%) attended 3 of the hub meetings, 18 (32.1%) attended 4 of the hub meetings, and 17 (30.4%) attended all 5 hub meetings. More than one teacher from a school may have attended a hub meeting. Schools sent at least one teacher to a median of 4, of the 5, hub meetings.

Table 15: ReflectED training attendance (intervention schools)

Training event (n = number of iterations)	Actual attendance, number of schools (% of 56 intervention schools)
Launch training Spring Term 2018 (n=5)	54 (96.4)
Regional hubs Summer Term 2018 (n=11)	46 (82.1)[a]
Regional hubs Autumn Term 2018 (n=7)	35 (62.5)
Regional hubs Spring Term 2019 (n=7)	39 (69.6)
Regional hubs Summer Term 2019 (n=7)	31 (55.4)

[a] Two schools were absent from this hub (not included in this number) and sent apologies, they had been in contact with the developers outside of the hub meeting.

ReflectED launch training and its cascade

Launch day training was attended by the headteacher and lead practitioner for most participating intervention schools, this training was then cascaded into each school by those who had attended the training at Rosendale Primary School by way of half a day inset training delivered in each school. The cascaded training delivered by each school was not observed, but questions about this were asked in both the self-reported training survey and the post-intervention teacher surveys for the intervention schools.

The launch day training covered key aspects of the ReflectED intervention as set out in the TIDieR (Table 1) and included demonstrations of key aspects including reflections on the Seesaw app. Each participant of the launch day training was provided with a printed booklet containing introductory materials related to the ReflectED launch training and the lesson plans for Term 1 of the intervention. Attendees were given a tour of Rosendale Primary School to see ReflectED in action, with opportunities to talk to staff and students about their experiences of ReflectED. Participants seemed particularly engaged with this and the opportunities to see ReflectED in action and discuss with teachers and students from Rosendale Primary School were well received. The observers (LG and LT) both noted the perceived value of seeing ReflectED in action at Rosendale Primary School, to complement understanding of the training and introduction to ReflectED that was being delivered. The training began with an introduction to theory that links to the ReflectED intervention, as developed by Rosendale Primary School. This aspect of the training included information and discussion about cooperative learning (e.g. Kagan), neuroplasticity, Growth Mindset, and metacognition. At the launch training there were questions from participating schools about whether ReflectED could be implemented without technology and this was confirmed, paper reflections were named as an alternative to Seesaw. However, there was emphasis that in this evaluation of ReflectED the aim was that participating schools would engage with Seesaw as the means of facilitating and recording student's reflections.

We received 69 responses to the online training survey that was distributed after the training events, of a total of 104 attendees thus a response rate of 66%, which should be considered when interpreting the results. The roles of the 69 training survey respondents were as follows:

- Headteacher: 39% (27 respondents);
- Deputy headteacher: 15% (10 respondents);
- Lead practitioner for ReflectED (in their school): 41% (28 respondents);
- Lead practitioner and head teacher: 2% (1 respondent); and
- Classroom teacher: 4% (3 respondents).

Of those respondents, 49% (34 respondents) did not engage directly in classroom teaching, 20% (14 respondents) were teaching in KS1 and 30% (21 respondents) in KS2. Four of the 69 respondents had previously engaged in ReflectED training, the detail given by them described attendance at conferences or training events or following the online ReflectED resources. Respondents to the training survey were asked to rate (on a 5-point Likert scale from 'strongly agree' to 'strongly disagree') the extent that they agreed that in relation to the training day:

a) they had the right amount of information about implementing ReflectED;

b) it was well organised;

c) the materials provided were useful; and

d) they were confident that they understood the principles of the ReflectED approach to metacognition.

There were 68 responses to these questions, with one missing response. The respondents were overwhelmingly positive about the training they had received, over 80% of the responses for all four points listed above were 'strongly agree' (a = 82%, 56 respondents; b = 88%, 60 respondents; c = 84%, 57 respondents); d = 90%, 61 respondents). Remaining responses were either 'somewhat agree' (a = 18%, 12 respondents; b = 10%, 7 respondents; c = 16%, 11 respondents; d = 9%, 6 respondents), or 'neutral' (neither agree nor disagree: b = 1%, d = 1%; 1 respondent each).

Around 96% (66 respondents) of the 69 training survey respondents reported that they then went on to train the staff in their school to deliver ReflectED, thus indicating that the respondents complied with the intended model of school staff trained by the ReflectED team subsequently delivering training to their colleagues when they returned to their own schools. The two respondents who said that they had not delivered training in their schools gave dates when they were going to implement this that were appropriate within the intervention timeline (i.e. it was not that they were not going to do it but rather they had not done it yet).

In relation to training staff in their schools, the training survey respondents were asked to rate (on a 5-point Likert scale from 'strongly agree' to 'strongly disagree') the extent that they agreed that:

a) the training day prepared them to deliver training to their colleagues back at school;

b) the ReflectED training was well received by teachers at their school; and

c) they felt the teachers in their school were prepared well to begin teaching using the ReflectED approach.

There were 64 responses to these statements (the 66 respondents who had stated they had already delivered the training in their schools were asked the questions and there were two missing). Again, the response was overwhelmingly positive with respondents predominantly strongly agreeing or agreeing. Other comments that respondents gave in the training survey were also largely positive, one headteacher commented: *'Excellent day—informative—all the information we needed to get the project off to a flying start.'* There were some comments around the amount of information conveyed during the training day being extensive, and recognition within this that some attendees had travelled quite far in one day to attend.

There were some suggestions that training might have been delivered regionally, however many of the free-text comments that respondents provided focused on how useful it was to see ReflectED in action at Rosendale Primary School in London. Semi-structured interviews supported the findings of the baseline survey relating to the training provided by Rosendale Primary School. Attendees reported that a lot of information was condensed into a 1-day training event, and while it was 'intense', overall, it was well organised, informative, and enjoyable. As the training was run by

those [Rosendale Primary School] who had already delivered/delivering ReflectED, teachers found the training to be very relatable to their own classroom and school and praised the quality of the information they received. Teachers found the observations of in-school metacognition lessons during the training day to be particularly valuable for understanding how ReflectED worked within the classroom. Following training, equipped with a bank of resources and strategies, teachers reported feeling confident to cascade the intervention to staff within their school.

> *It [the training] was overwhelming, quite overwhelming, but it was useful watching lessons and watching how it was being delivered in a school that's obviously got it quite well embedded. But yeah it was a lot to take on in that day. I think it could have done with being a bit of a longer training, maybe two days, maybe having the training down there and then them coming into school, because it was an awful lot to take in. (Classroom teacher, intervention school)*

> *… really nice to be able to go to [the training], for us anyway, it was a very contrasting school, and see some of the sessions being observed, I mean, in an ideal world it would have been nice to spend, like, a whole day observing different reflective lessons in the different classes. But, no it was really good, very well organised and, yeah, prepared us for what we needed really. (Classroom teacher, intervention school)*

> *… it [the training] was quite relatable, seeing the teaching staff speaking about their experience, because you can sort of think, oh well my class could do that. It wasn't just a stand alone training and we expect…it was relatable because they'd done it as well. (Classroom teacher, intervention school)*

Those who attended the training reported cascading it to others within the school. As part of the cascade training staff recalled completing the ReflectED activities, such as juggling, which they found not only to be useful practice for delivering the lessons, but also to aid their understanding of how the learning from the lessons may be perceived from a child's point of view. Overall, the cascaded training was well received by teachers. This, coupled with the '*very detailed*' *(classroom teacher, intervention school)* lesson plans, meant staff felt sufficiently prepared to deliver the lessons within their own classrooms.

> *I think to a point because we got to try out the different activities but from, obviously, lower down all the way up to the top. And then we, sort of, talked about how it felt for us because a lot of people couldn't actually do some of the things, like the juggling and stuff. So, you could see, actually take it down to the children's level, how they find it very difficult. And then, obviously, they explained the process of how it works through the programme, and by the time they get to the end this is what they should be able to do. It was really helpful trying the activities out first before we actually taught them. (Classroom teacher, intervention school)*

> *It was really useful. It was really useful. I mean, it was done by [name] and our assistant principal. They'd both done the training, and they shared, really, I think, almost everything that they'd learnt down there. And I think they did a very good job of selling it to us. (Classroom teacher, intervention school)*

> *It was so explicit in the teaching handbooks which are fabulous, very little direct training was required. (Classroom teacher, intervention school)*

> *We received, I think it was one or two staff meetings by someone that had been on the training. We were shown very clearly how to use Seesaw, how to access all the plans and resources online. The thing that we were shown how to use most was the Seesaw because it was something we hadn't used before. But then all of the plans and resources and everything were so self-explanatory, it was really easy to implement quite quickly. (Classroom teacher, intervention school)*

One teacher discussed the importance of continuing to receive cascade training for retaining the momentum ReflectED had gained within the school.

> *I think it would be handy to, kind of, revisit it […] and refresh the training. If that doesn't happen I think we're in danger of, because we're going, we have a lot of changes, as always, in education. I think something else would come in and we'd kind of go, oh actually we've been doing Metacognition for a while, let's just let that drop off for this week, and then after it drops off for one week that's it. So, I think,*

if [name] just kept refreshing it for us and, you know, assuring us that it was, reassuring us that it is a priority in school. But we need that, I think, from SLT [senior leadership team] really, don't we, to come down to us. (Classroom teacher, intervention school)

Some teachers described a 'whole-school approach' to the cascade training that involved disseminating the training to all school staff and sharing the approach to children and their parents/carers. On a staff level, this involved staff meetings to disseminate ReflectED to support staff as well as teachers. Teachers also described holding assemblies to introduce the schoolchildren to ReflectED and one school explained that they had communicated to parents about ReflectED through a blog (Classroom teacher, intervention school).

The teachers that went on ReflectED, they came back with a bank of resources and strategies and ideas to introduce to the rest of the staff within the school. Everyone, not just teachers, but teaching assistants and everyone that works in the school. (Classroom teacher, intervention school)

It was just a case of coming back, then, and digesting everything before I did the [cascade] training; and then I fed the training to staff, initially, teaching staff, initially. And then when I introduced it to the children, the teaching assistants came to the assembly when I introduced it to the children [via an assembly], so that they had a knowledge, as well. (Classroom teacher, intervention school)

In the post-intervention survey administered to Year 2 and Year 6 teachers, headteachers and nominated leads in the intervention group (n=179) several questions were asked in relation to training received. These questions sought to understand the cascade model of training. Many teachers received their training from other sources instead of, or as well as, the cascaded model within schools as detailed below (more than one response could be given).

- 46% of respondents (or 82) had attended the initial launch training at Rosendale Primary School between January 2018 and February 2018, which reflects that the survey was targeted at nominated leads and headteachers who were expected to attend the Rosendale Primary School training as well as at classroom teachers.
- Over half of all respondents (54%, n=96) reported that they received training in their school delivered by teacher(s) from their school in Spring Term 2018 although four respondents also indicated that they received training in Autumn Term 2017 and a further two referred to 'Ongoing training in staff meetings' and 'Additional updates provided by the school head'.
    - 24% (42 respondents) reported that they had been trained informally by colleagues at their school (e.g. talking through what ReflectED is about, watching lessons, but not a formal CPD presentation in school);
    - 9% (16) reported that they received training at the lead school for their regional hub with the team from Rosendale Primary School in September 2018;
    - 3% (6) reported that their school had a visit from the ReflectED team (from Rosendale Primary School) and they received or had training during that visit; and
    - only two respondents indicated that they had not received any training for ReflectED, one explained that they had used the booklet for lessons and the other that a previous staff member had received the training but had left the school.

In relation to the last point in the list above—in the developer interview at the end of the evaluation period, it was confirmed that additional training for school staff who were perhaps new or had missed the initial training at their school was offered by Rosendale Primary School (in the 2018/2019 academic year of the intervention), however they reported that the uptake of this was not high.

All respondents who indicated that they had received ReflectED training (n=177) were asked their agreement with a series of Likert scale statements as presented in Table 18. Most respondents were positive about the training received; notable findings included:

- 97% (174 out of 177) agreed (strongly or somewhat) that the training they received for ReflectED was pitched at an appropriate level;
- 96% (173) agreed (strongly or somewhat) that the training they received prepared them for their involvement in the delivery of ReflectED in their school; and
- 97% (173) agreed (strongly or somewhat) that after training they were confident that they understood the principles of the ReflectED approach to metacognition.

Only one respondent indicated that they somewhat disagreed that the training prepared them, the same respondent also indicated that they somewhat disagreed that they felt confident that they understood the principles of the ReflectED approach after training.

Table 16: Intervention respondents' views on ReflectED training they received (n=177)

Statements	How far respondents agreed: number (%)				
	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
The training I received for ReflectED was pitched at an appropriate level	133 (74.3)	41 (22.9)	3 (1.7)	0 (0.0)	0 (0.0)
The training I received prepared me for my involvement in the delivery of ReflectED in my school	123 (68.2)	50 (27.9)	3 (1.7)	1 (0.6)	0 (0.0)
After training I was confident that I understood the principles of the ReflectED approach to metacognition	116 (64.8)	57 (31.8)	3 (1.7)	1 (0.6)	0 (0.0)

Nominated lead teachers were asked about the support they felt they received from Rosendale Primary School, their confidence in delivering training within their school, and the responsiveness of colleagues to the training and ReflectED. As can be seen in Table 19:

- 96% (47 out of 49 respondents) agreed (strongly or somewhat; 42 and 5, respectively) that if they had a question about ReflectED it was easy to contact the team at Rosendale Primary School and get help;
- 94% (46 respondents) agreed (strongly or somewhat; 35 and 11, respectively) that they were confident to deliver guidance and training to other staff in their school about ReflectED;
- 90% (44) agreed (strongly or somewhat; 25 and 19, respectively) that the training they delivered about ReflectED in school was well received by staff; and
- 88% (43) agreed (strongly or somewhat; 21 and 22 respondents, respectively) that staff in their school were enthusiastic about ReflectED.

Table 17: Nominated leads reponses relating to delivering training and receiving support (n=49)

Statement	How far respondents agreed: number (%)				
	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
If I had a question about ReflectED, it was easy to contact the team at Rosendale and get help	42 (85.7)	5 (10.2)	2 (4.1)	0 (0.0)	0 (0.0)
I was confident to deliver guidance and training to other staff in my school about ReflectED	35 (71.4)	11 (22.5)	2 (4.1)	1 (2.0)	0 (0.0)

The training I delivered about ReflectED in school was well received by staff	25 (51.0)	19 (38.8)	4 (8.2)	1 (2.0)	0 (0.0)
Staff in my school were enthusiastic about ReflectED	21 (42.9)	22 (44.9)	5 (10.2)	1 (2.0)	0 (0.0)

Regional hub meetings

The regional hub meetings that were observed by LG and LT were found to be supportive places for schools participating in the ReflectED intervention to come together, with other schools in their geographical region, to share good practice with each other and to receive focused input from the ReflectED team. The findings of observations of five regional hub meetings held between May 2018 and March 2019 were in line with the stated purpose of regional hubs for 'training and troubleshooting' (see Figure 1) and as well as formalised input from the developers (e.g. distribution of lesson plans and discussion of them for forthcoming terms) there were more informal opportunities for teachers to discuss and explore how they were implementing ReflectED in their school settings.

Nominated lead practitioners were expected, as per the MOU, to attend termly regional hub meetings. Around 38% of respondents (67 out of the 177 who responded to this question) indicated that they had attended regional hub meetings. The average number of hub meetings attended was 3 (attended by 31% of respondents, or 21 out of 67). Around 19% (13 respondents) attended only one hub meeting, 19% (13 respondents) attended two hub meetings, 15% (10 respondents) reported attending four hub meetings, and 15% (10 respondents) reported attending five hub meetings. These 67 respondents were also asked to rate their level of agreement with several statements relating to the hub meetings. The results can be seen in Table 20 and demonstrate high levels of agreement in relation to perceived usefulness of the regional hub meetings, although perhaps less strongly that those found for the ReflectED launch training.

- 89% (60 out of 67 respondents) agreed (strongly or somewhat; 38 and 22 respondents, respectively) that the hub meetings were useful because they could share practice;
- 88% (59 respondents) agreed (strongly or somewhat; 38 and 21, respectively) that they picked up useful tips to implement ReflectED at the hub meetings;
- 97% (65 respondents) agreed (strongly or somewhat; 47 and 18, respectively) that they were able to have any queries they had about ReflectED answered at the hub meetings; and
- 91% (61 respondents) agreed (strongly or somewhat; 42 and 19, respectively) that after attending hub meetings they communicated the discussions to their colleagues back at school.

Table 18: Intervention respondent views on the hub meetings they attended (n=67)

	How far respondents agreed: number (%)				
Statement	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
The hub meetings were useful because we could share practice	38 (56.7)	22 (32.8)	6 (9.0)	0 (0.0)	1 (1.5)
At the hub meetings I picked up useful tips to implement ReflectED	38 (56.7)	21 (31.3)	5 (7.5)	3 (4.8)	0 (0.0)
I was able to have any queries I had about ReflectED answered at hub meetings	47 (70.2)	18 (26.9)	2 (3.0)	0 (0.0)	0 (0.0)

After attending hub meetings, I communicated the discussions to colleagues back at school	42 (62.7)	19 (28.4)	5 (7.5)	1 (1.5)	0 (0.0)

Headteachers (n=48) were also asked separately about staff attendance at hub meetings and support from Rosendale Primary School. Around 75% of headteachers (n=36) agreed (either strongly or somewhat: 40% and 35%, 19 and 17 respondents, respectively), 10% (n=5) neither agreed nor disagreed, and 15% (n=7) disagreed (either somewhat or strongly, 10% and 4%, 5 and 2 respondents, respectively) that they were easily able to free up staff from school to attend regional hub meetings (note: numbers may not add up due to rounding). Of those 12 respondents who neither agreed nor disagreed or who disagreed with this statement, cited the pressures and cost of releasing staff, alongside geographical distance needing to be travelled, and an increased workload for colleagues.

Staff who attended training hubs spoke highly of them during the semi-structured interviews; one attendee described them as being more useful than the initial training event. Training hubs presented an opportunity for teachers to network with other schools and share practice relating to the delivery of ReflectED within and across their schools.

> *I think the hubs have been more useful because you're talking about it on a level at the same level as you are. So being able to go and say, we've tried this, this hasn't worked, this has and … It was obviously useful to go and see a school where they've had it but I think because of the distance and everything else as well, the hubs have been much more useful because you're talking to people that are at the same level as you are. And when you're talking about, have you developed a policy or have you done anything around that, or, we've tried this out … and you get lots of different ideas, things like that. So, we've enjoyed the hubs. (Classroom teacher, intervention school)*

> *There's lots of things that we're gathering up when we go to the Hubs, loads of ideas that … we would like to do. (Classroom teacher, intervention school)*

85% of headteachers (41 out of 48 respondents) also agreed (73% strongly and 13% somewhat; 35 and 6, respectively) that if they had a question about ReflectED, it was easy to contact the team at Rosendale Primary School and get help. A further 15% (7 respondents) neither agreed nor disagreed with the statement (note: numbers may not add up to 100 due to rounding). Data gathered for the IPE suggests that the developers provided a good level of support that maximised the chances of compliance and fidelity to intervention as far as possible.

Fidelity

Data collected in relation to fidelity aims to facilitate a description and analysis of the extent to which the ReflectED intervention was delivered as intended in the intervention schools. Fidelity will also explore any issues with fidelity to the ReflectED intervention as intended and the reasons for these issues.

Data sources:

- structured observations;
- teacher interviews;
- pupil focus groups; and
- post-intervention survey (intervention schools).

Research question(s):

8. To what extent do the schools and teachers implementing ReflectED adhere to the intended model of ReflectED as a whole-school intervention?

8b. What variability in the implementation of ReflectED exists across different participating settings? Are there any barriers?

9a. How well have different components of ReflectED been delivered and how well have participants engaged with it?

11. What is the reach of ReflectED across the intervention schools? … Are there any perceived or actual benefits for specific groups of pupils (e.g. SEN, EAL, GRT)?

12. Has ReflectED been adapted in any way during the intervention period? In what ways and why?

13. What evidence is there, in pupils' reflections, of metacognition and any change in this over the course of the intervention?

ReflectED as a whole-school intervention (research question 8)

Findings summary:

ReflectED was intended to be delivered, in this evaluation, as whole school intervention with delivery across the primary school age range of 4–11 years (see Table 1) and across different subjects within the curriculum. The data included in relation to research question 8 are teacher interviews, pupil focus groups, and the post-intervention survey. Analysis of this data supports that ReflectED was being delivered as the intended whole-school intervention, both in terms of delivery across age groups and associated classes and within and between different subjects in the curriculum (i.e. not limited only to the weekly ReflectED lesson). The responses gathered from teachers did not make any reference to GRT (cf. research question 11), so there is no reference to this in what follows.

Data from teacher interviews and pupils focus groups provided examples of schools implementing ReflectED as a whole-school intervention. Such evidence presented in two areas: i) the delivery of ReflectED across the school, and ii) the transfer of metacognitive skills (the ReflectED approach) into other subjects. For example, one pupil shared that their school had adopted badges to identify pupils who were 'metacognition masters'. Here, children across the school could approach those wearing the badges and ask them for help when they were struggling with a task.

> *[the badge] indicates that children around school know that they can come to us or come to anybody say and then talk to us about stuff that they're struggling on and they need work and then visitors know that we're the people who they can look out for and talk to and if they need help they can talk to us. (45017, Pupil 1)*

During the interviews, teachers often discussed how ReflectED was being implemented across different year groups within the school. For example,

> *We've used quite a few of the templates. We've also, we've set up in Maths in particular, Key Stage 1 anyway, like a pre-set template that our learning objective goes on so they've always got that colour reflection at the top anyway. As you go up to year six there's been a bit more free-flow, kind of, writing it on the bottom of the book. But we introduced the templates that were introduced to us at lots of the hubs and things and online, and then we have RelfectED books as well. So, either … if it's cross-curricular, then it's obviously in those books, it's in the Maths books, it's in the literature books, but if it's been a specific ReflectED lesson, then we've got a ReflectED book to be able to put that evidence from that lesson in. We're monitoring tomorrow. (Classroom teacher, intervention school)*

Teachers related the success of implementing ReflectED across the school to the structured lesson plans (the usefulness of the ReflectED lessons plans is also detailed in the 'Training' section in relation to research question 8a).

> *So, I think, you know, in terms of getting it across such a large school it's gone quite well, particularly for that reason [the lesson plans]. Because it's there, it's planned, it's a set structure, it's a set way of working and everyone's working it the same way so when we have staff meetings – we had one a couple of weeks ago – come and share practice, tell us what you've done, people could bring things and then share ideas as well and they can see how that can fit in cross-curricular for them as well. So yeah, no. Just using it. (Classroom teacher, intervention school)*

But other than that, I think, the lessons are … they're planned out for you, that's made it really easy to be able to introduce it across school and say look, it's here, it's not that you've got to think about it and it's developed teacher understanding as well because of how they're already set out for them. (Classroom teacher, intervention school)

In addition to the weekly *ReflectED* lesson, it was also evident within some schools that teachers were transferring the *ReflectED* approach to learning in other subjects. One teacher stated they incorporated *ReflectED* into *'pretty much every subject' (Classroom teacher, intervention school)*, and pupils also noticed the transfer for *ReflectED* into other subjects.

I use it in maths. I would use it in science. We've used it, for example, we did editing skills in writing so we used it then, and if it's kind of a different learning experience then I would use it then. (Classroom teacher, intervention school)

We don't just use them in the ReflectED lessons because sometimes when we're in a lesson the teachers will ask you what colour you feel you are at that moment, and when [the teacher] comes around, she always talks about it and gets you to talk about the colours. (Pupil, intervention school)

[We use ReflectED in] most of the lessons, [teacher's name], when we do some of the lessons, [teacher's name] always asks us how we felt on that lesson and if you think you've improved or if you're still there. (Pupil, intervention school)

Most often teachers reported using ReflectED in core subjects like mathematics and English, and other subjects such as Physical Education (PE) and Science. The ease of delivering ReflectED into other subject/lessons varied, with many considering it easier to transfer to subjects/lessons that have an objective outcome. For example, during interviews many teachers and pupils discussed that it was easier to use ReflectED in mathematics than it was in English.

Maths tends to be more, kind of, mechanical so, I don't think I can do this, or I think I can, or I've done this, or I don't get it yet. Whereas English is more, I suppose, it's more reflective in that the children are actually reflecting on the content of what they've done, because English lends itself more to that, they work in both ways, they're just different. (Classroom teacher, intervention school)

You know, Maths in particular, they reflect at the start of the topics, they know where they were, and they find it very easy to say, well I'm now at this point at the end of the topic and this is the reason why. English I think they do find harder because they're having to do so many more things in order to get there, but that's something that, you know, we're working with them. (Classroom teacher, intervention school)

Yes, by incorporating it in like every subject as well. So for example when we started PE we were doing ball skills and then I showed them a demonstration of what I wanted to do. I said right, assess your skills, where do you think you'll be and we just put the colour codes out. So it was kind of like trying to gain as much as possible. (Classroom teacher, intervention school)

Yes, within ... it's easier within maths at the moment, because the strategies are easily broken down and they can say, yes, I could do this, no I need to work on this and they're able to identify what helped them or what could help them next time. English is a bit more tricky but they're getting there. (Classroom teacher, intervention school)

Interviews with teachers suggested that ReflectED was complementary to the curriculum, with one school noting that ReflectED had supported a whole cross-curricula approach to learning.

So, it is good because that's then helped us with that whole cross-curricular approach. So that's what we were trying to say to staff, put it in other lessons. But they still saw, this is ReflectED, and I can do a bit of reflection in these. Whereas because you've got to take the lesson and put it into science or for the planning or something like that. It almost forces you to think about where you can fit it in and where you can do it so I can put the plan one in the DT [Design and Technology] lesson or I could … so that third one is quite nice in the way that it takes you off and helps you to develop it more through other lessons. (Classroom teacher, intervention school)

It's worked really well across the board, and it's, really, it's complemented a lot of what we do anyway. (Classroom teacher, intervention school)

Teachers also discussed how pupils were transferring the metacognitive skills and tools used within ReflectED to other subjects/lessons and we explore the data relating to this in research question13. In the post-intervention survey, Year 2 and Year 6 teachers (n=97 respondents) were asked about the extent to which they could see pupils applying things they had learned in ReflectED lessons in other lessons—90% (87) agreed (41% strongly and 49% somewhat; n=40 and 47 respondents, respectively) that they could see pupils applying things they had learned in ReflectED lessons to other lessons, and the remaining 10% neither agreed nor disagreed.

In the post-intervention survey headteachers (n=48) were asked about their level of agreement with several statements surrounding ReflectED as a whole-school intervention. There was slightly less agreement surrounding the ease of fitting ReflectED into the school curriculum, although a still high 85% of headteachers (41 respondents) reported strongly (50% or 24) or somewhat (35% or 17) agreeing. 6% (n = 3) neither agreed nor disagreed, 6% (n = 3) somewhat disagreed, and 2% (1 respondent) strongly disagreed. This level of agreement was also noticeably higher than the 63% of Year 2 and Year 6 teachers who agreed that the intervention was easy to schedule into their weekly plans/lesson schedule reported above (see Compliance section). However, 96% of Year 2 and Year 6 teachers (94 out of 98 respondents) did agree (60% strongly and 36% somewhat, n=59 and 35, respectively) with the statement 'I have used aspects of ReflectED in my other lessons'. The remaining 4% neither agreed nor disagreed. Those teachers who did agree referred to using reflection sheets, and Seesaw and, most frequently, the performance tag colours (81 respondents out of 93 respondents to this question). The Year 2 and Year 6 teachers also mentioned using aspects of ReflectED across the curriculum although especial mentioned tended to be focused on literacy and mathematics.

Variability in implementation, barriers to implementation and adaptations (research questions 8b and 12)

Findings summary:

This section explores variability in the implementation of ReflectED, during the intervention period, and any barriers and adaptations to implementation. The data presented here is from the post-intervention survey, teacher interviews and the developer interview. The main example of variability identified was around how reflections were completed and recorded, from early in the evaluation period it became apparent that schools were struggling to record reflections of the expected rate per pupil on Seesaw. This was mitigated for early in the evaluation period by the developers and alternatives (including paper-based reflection templates) were provided. The main barriers that the IPE data illustrates were linked to the variabilities in implementation and focused on the completion of reflections using Seesaw and the time taken to teach a ReflectED lesson (over and above the 30–40-minute allocated time in some cases). Schools and individual teachers made relatively few adaptations to the ReflectED intervention, which aligns with the compliance data presented in that section of the report.

Semi-structured interviews with teachers identified variability in how ReflectED was implemented across schools. The interview data revealed that the variability was driven by barriers to implementation. The key themes to emerge from the interview data included: i) how pupils completed and recorded their reflections; and ii) the timetabling of ReflectED lessons into the curriculum. There was a great variability as to how and when pupils completed and recorded their reflections, which deviated from the ReflectED 'gold standard' of recording reflections to Seesaw. Teachers reported completing reflections verbally (discussions), on paper (sometimes transferred onto Seesaw via an iPad by the pupil or the teacher), or in children's workbooks.

Yes, I mean, at upper Key Stage Two, sometimes with it being SATs [Standard Assessments Tests] led, we do sometimes struggle to get it in, but we do ensure that it is timetabled in and we are teaching it rather than using it as kind of an accessory, we use it as an integral part as well, so during the lesson, it's not always write yourself an observation, it's kind of tell us how you feel about how you've learned, what have you learned that's helped you to learn and to achieve what you needed to get to, so it's more interwoven rather than a discrete lesson at times, but it's still at the very forefront of their minds when they're doing certain types of lessons. (Classroom teacher, intervention school)

Teachers of younger children (KS1 and younger) seemed to make choices to complete the reflections on paper with their classes as the technology (accessing Seesaw using an iPad) was too complicated for them. This sometimes

resulted in an increased workload for teachers as some would upload the paper reflections to Seesaw, which they viewed as burdensome.

> *I think that needs some consideration because the children are too little to do it themselves. (Classroom teacher, intervention school)*

> *Yeah, I think once the children, particularly in early years, are more adept at using it themselves it'll be less of a burden on the teachers. (Classroom teacher, intervention school)*

> *I would say for this age using the iPads at this point to get them to do it on their own, I think as it gets later in the year, I'll find it easier but at this point, obviously you [inaudible] how disastrous it can be, trying to get them all to do it. But eventually when they get it, it's fine, but at this point, they're having to get onto obviously using the QR [Quick Response] code which they're getting better at using QR codes, but getting it then putting it in their own file and just getting them to do it independently. I think as they got a bit older, I think it would be easier for them to just go, yes, I'll go get the iPad, I'll put it on and do it myself, where I'm just constantly checking that everything's … we're overseeing everything, if that makes sense? (Classroom teacher, intervention school)*

Teachers across several schools noted that it was easier for KS2 pupils to gain access to iPads to complete reflections via Seesaw, however many schools were unable to use Seesaw as they did not have access to iPads. Survey and interview data highlighted that there were a range of reasons for this. Post-intervention survey data asked Year 2 and Year 6 teachers who reported not using Seesaw (n=70) the reasons for not using Seesaw to reflect. Around 51% of teacher respondents reported that there were not enough iPads in school, 36% reported that the time taken to use Seesaw was too long, 31% reported that pupils were not able to operate Seesaw with some degree of independence, 19% cited internet speed as problematic, and 34% of teacher respondents gave other reasons including that paper-based reflections were more efficient, and limited sets of iPads (e.g. only one set in the whole school).

Interview data illustrated that in most cases, schools only had a limited number of central iPads, which needed to be reserved for a class to use; timing the access to iPads alongside the ReflectED lesson was not always possible. Often a small number of iPads needed to be shared between a whole class, which was logistically challenging for class teachers during lesson time. While some teachers preserved, others uploaded paper reflection to Seesaw outside of lesson time that often restricted the number of reflections that were uploaded to Seesaw or refrained from engaging with Seesaw completely.

> *We gave up on that [uploading reflections to Seesaw] really early. Yeah, I think just because of the size of the school, particularly with Seesaw, being able to actually get that and upload it and it wasn't as … it's not a, you've got your iPad there and then and you can just go with it. (Classroom teacher, intervention school)*

> *I know Rosendale are very lucky to have a lot of technology, compared to [our school] … we don't have that and we have to timetable our iPads out once a week for each class to reflect. So, I think the children love it and if they had the opportunity to spend all day on Seesaw, I think they would … when you've only got sort of six iPads between, you know, 30 five and six year olds, it is really demanding. It is hard. But we've had a good go. (Classroom teacher, intervention school)*

> *I would say the thing that's most difficult, it's probably a question later on, is we really struggled with the Seesaw element massively, because of a lack of just technology in school and equipment. Sometimes we have like one or two iPads per year group, and we just don't have any money in the budget at the minute to buy any more. (Classroom teacher, intervention school)*

> *If we're going to do it way that Rosendale want to do it they need to make sure they've got resources to use Seesaw because of the fact that everything's had to be uploaded and I don't know if that's just for data collection purposes, but you can do it completely without technology. It just depends how you want to kind of put your reflections. Because when we looked at it originally, and I didn't want all the sheets stuck in the books. We didn't want a whole book full of reflection sheets. It just depends I suppose how schools want to go about doing it. They like doing it in their own way. Because we haven't ... like year*

six, we've not used a lot of those sheets at all, they've just done their own thing in the book, but then lower down you know. (Classroom teacher, intervention school)

In the 16 lesson observations that were completed, across the evaluation period, of the 16 lessons 7 included the use of Seesaw to reflect, 9 included the use of reflection sheets to reflect (with 3 observations showing simultaneous use of Seesaw and reflection sheets seemingly to mitigate for access to iPads when there were not enough for the whole class). In 2 of the 16 observations there was not a specific time for reflection observed. The developer interview illustrated how the development team were aware from an early stage in the evaluation that there were some issues with intervention schools having difficulties in completing reflections on Seesaw and alternatives were suggested and communicated to schools in the weekly emails and via the regional hubs:

… about a term and a half in a lot of the schools were struggling with getting it down [reflections on Seesaw] as much as we had anticipated they would onto Seesaw. So, what we did was that we said that they had options: they could just choose two children to do Seesaw with or they could do it with a whole class … what we realised was that we were in danger of losing a number of schools because of what they saw as the commitment to Seesaw. (Developer interview)

And there was clarity that the technology (Seesaw) was intended to facilitate the process of reflecting rather than it being the process of reflecting itself:

And additional technology is there to facilitate the process. The process isn't dependent upon the digital technology, but if you can get it working it does make it a lot simpler to do. (Developer interview)

Interview data also highlighted variability in how schools found time to complete the weekly 20–30-minute ReflectED lessons. Most schools reported ringfencing time to provide a fixed timetable slot for ReflectED lesson, with an aim of embedding metacognition into the curriculum. Teachers reported that staff within their schools' raised concerns about how ReflectED lessons could be squeezed into their class timetable. In all cases, incorporating ReflectED into the timetable came at the expense of other themed subject lessons, e.g. Personal, Social, Health, and Economic, PE, Music, and Languages; core subjects did not appear to be affected by such timetabling changes.

It has been tricky to fit it in, but for me, it's been sticking to a timetable. If it's as and when, I don't think it works. It has to be a slot every week where you really dedicate that time, because, you know, by Thursday you've got 101 jobs and you think, oh. I just think it's got to be a slot, like your mathematics, your English, your topic. It's got to be that dedicated timeframe. For me, that's what works for me, so that's what we've done in year one. (Classroom teacher, intervention school)

Every year group timetables it when it's convenient for them. I think it is hard because we do have a very busy curriculum as it is and it's very full, and it's like, you know, you've got to fit in ten minutes of handwriting a day, ten minutes of this a day. Spelling and all that kind of stuff and obviously it taking half an hour is ... a lot of our subjects are like an hour, so then it's like, right, well what are we going to do for that other half an hour that complements it, and sometimes that's been a bit difficult, but everyone has managed it. There might have been a few moans and groans, but it is kind of one of those things where people just accept it now, so at the beginning it was a bit like, oh, I can't fit it in. But, now everyone knows metacognition fits in and it's part of the curriculum so they just to do it and people don't say anything about it now, they just do it. So, that's quite good. (Classroom teacher, intervention school)

Teachers discussed that timetabling ReflectED lessons was more difficult for older pupils (KS2) in comparison to KS1 pupils due to curriculum demands.

We've made space. We've had to just make space. It's harder in year six because we've got so many other things and we've got something called reading plus as well which you have to find a standalone lesson for reading in the afternoon. There's just so many things. But we've made it, we've made it fit. (Classroom teacher, intervention school)

Another aspect of the ReflectED intervention where schools reported variability was in pupils working in mixed ability pairs. Around 59% of Year 2 and Year 6 teachers (57 out of 97 respondents) indicated that their pupils always worked in mixed ability pairs during ReflectED lessons and a further 32% (31 respondents) indicated that this occurred most of

the time. This is supported by structured observation data—of the 17 ReflectED lessons observed by the evaluators, in 16 evidence 'often' or 'sometimes' was observed within the observation period of children being given 'plenty of opportunity' to share ideas with their partner. However, 5% (5 respondents) indicated that this occurred only half of the time and 4% (4 respondents) only sometimes. Of the 26% of Year 2 and Year 6 teachers (25 out of 98 respondents) who indicated that their class was a mixture of two year groups there was a mixture of approaches, with the year group lessons taught often dependant on the perceived ability of the pupils, sometimes involving a 'pick and mix' approach to choosing, which lessons to teach by the teachers or alternating the year group lessons taught by year (i.e. Year 5 in one year and Year 6 lessons the subsequent year with the same group).

There was much less variability, and more compliance, with the use of the performance tag colours as reported by teachers in the intervention schools in the post-intervention survey: 90% of Year 2 and Year 6 teachers (88 out of 98) agreed (64% strongly and 26% somewhat; 63 and 25, respectively) that they used the performance tag colours with their class regularly; 4% (4 respondents) neither agreed nor disagreed; 5% (5 respondents) somewhat disagreed; and 1% (1 respondent) strongly disagreed.

In semi-structured interviews, teachers discussed how the 20–30 minutes allocated to a ReflectED lesson was sometimes too short and often a barrier to completing the required activities within the allocated time. For teachers of younger children, this was due to them having to provide extended explanations/instructions, the children taking longer to complete the skills activities and/or to reflect. For older cohorts, this was because they were required to complete more detailed reflections and more discussion at times. Despite this, teachers saw value in the lessons and attempted to accommodate where possible; some teachers ran-over their timetabled slots, some cut the lesson short, or chose not to complete the reflections. Teachers considered that ReflectED lesson time slot could be extended to 40–60 minutes. Teachers who had delivered ReflectED to a previous cohort noted that their own experience of the lessons meant they could manage and plan their lesson time better.

> *So there have been certain lessons that they've definitely taken at least forty-five minutes … there was one on building blocks where we were doing what makes a good learner and a bad learner. And one of the tasks was to use a cube base but then try to build the highest block. So that one, it took an hour. And then we had to obviously take the pictures and put it in the folders and stuff. But, for example, that one I didn't mind going over a little bit because it was engaging and there were so many different skills that they were learning from it, it was fine. (Classroom teacher, intervention school)*

> *I think that because I've repeated certain lessons, I think that's made it easier because I'm already aware of what needs to be done but sometimes it can … the only thing I would say implementation-wise is sometimes it takes longer than the twenty minutes because you have to give a lot … especially in the early years. (Classroom teacher, intervention school)*

> *Sometimes it is difficult, sometimes if it's a more intricate activity, I'll just take longer and that's fine, but you can kind of judge it from looking at the plan in advance. (Classroom teacher, intervention school)*

> *I think the lessons themselves, they're really good but there's a lot in and quite often I don't get enough, you know, and especially with that class, because they get into a discussion or we get into a … and then you think … oh, and it's just one more thing that I think, oh my gosh, I'm behind, you know? But I honestly don't think it matters as long as you're doing something like that on a regular basis because it then keeps everything fresh in their heads. (Classroom* teacher*, intervention school)*

Year 2 and Year 6 teachers in intervention schools were asked in the post-intervention survey about the extent to which they felt that ReflectED lessons could be taught in the suggested 20–40 minutes. The majority (79%) of teachers who responded to this question (77 out of 98 respondents) agreed that 'ReflectED lessons can be taught in the suggested 20–40 minutes': 41% (40 respondents) strongly and 38% (37 respondents) somewhat agreed; 8% (8 respondents) neither agreed nor disagreed; and the remaining 12% (13 respondents) somewhat disagreed. The respondents who somewhat disagreed (13 respondents) were subsequently asked how long, on average it took them to teach a ReflectED lesson. Across the 12 teachers who answered this question, the length of time varied from 45 minutes to 60 minutes with an average across the sample of 52 minutes a lesson. When asked why they felt this to be the case over half (7 respondents) mentioned the reflections, both giving time for pupils to reflect on their learning and recording their reflections (e.g. on Seesaw). Four teachers mentioned the time it took children to discuss in ReflectED lessons, often also mentioning that they saw this as an important aspect of the programme that they did not feel should be missed.

Three teachers indicated that the concepts were complex, and it took time for children to grasp them and a further two teachers identified the skills lessons as being particularly time consuming.

During focus groups, pupils also highlighted that they struggled to complete the required activities during ReflectED lessons as there was a lot to get through in a short space of time.

> *I think one of the things in the actual ReflectED lessons is that when the teacher is at the front explaining everything and sometimes they have to do it quickly to fit it into the time, sometimes it can be hard to take it all in and sometimes I'll be left like, what, but yeah, maybe have a bit more time to explain. That's about it I think. (Pupil, intervention school)*

> *I think we should have longer on the things that we're doing because we've just done the compass and I don't think we've had enough time yet. (Pupil, intervention school)*

Linked to the time taken to complete ReflectED tasks and activities, teachers also identified that the complexity of the ReflectED tasks/activities was a barrier to some groups of children (e.g. younger cohorts, lower ability children, disadvantaged children, and children with SEN and Disabilities [SEND]) engaging with the lessons. This is discussed in detail in relation to research questions 9 'engagement' and 11 'reach' of the ReflectED intervention.

Relatively few adaptations were reported by intervention schools as having been necessary to implement ReflectED. In the post-intervention survey, headteachers and Year 2 and Year 6 teachers were asked if the intervention was adapted to fit with any existing approaches (to metacognition and otherwise) in the school. Of the 177 respondents to this question: 42% (75 respondents) strongly agreed; 35% (62 respondents) somewhat agreed that such adaptations had been made; 4% (7 respondents) somewhat disagreed; and 2% (3 respondents) strongly disagreed. A further 17% (30 respondents) neither agreed nor disagreed. The adaptations ranged from adapting the school culture and curriculum to incorporate the ReflectED approach to tailoring the lessons to suit pupils needs (e.g. in choosing ReflectED lesson plans that were not necessarily aligned with the year group taught). Adaptations were explored in more depth in the teacher interviews, where the main adaptations focused on reflections and the format in which they were recorded (e.g. use of paper reflections, notes in workbooks, or discussions as opposed to Seesaw), and the frequency in which they were completed, which we presented above. One school reported a significant adaptation, which was that the HLTAs delivered the ReflectED lessons instead of class teachers (the intended method of delivery).

> *I mean we have our HLTAs; they deliver like specific metacognition in our PPA [planning, preparation, and assessment] tasks. So, they have a discrete lesson there. So, we don't teach it as a discrete unit ourselves as class teachers but what we do do is we filter out the aims of it in our lessons. So, we don't teach it explicitly like that. (Classroom teacher, intervention school)*

Aside from these points above, teachers reported very few other adaptations during interviews. Teachers were asked whether they had adapted the ReflectED lesson plans. Most teachers responded that they had not. This was mainly due to lesson plans being '*explicit*' *(Classroom teacher, intervention school),* making them easy to translate and deliver within the classroom without major adaptation.

> *The lesson kind of plans that we've been following, they're really helpful. I can't think of anything that we'd change about them because you just want a snapshot and then to be able to go off and do it and that feels like what that is. (Classroom teacher, intervention school)*

> *No, I wouldn't have said so, no, I just try and do what it says because it's a Noddy's guide, really, isn't it? But I've looked at one or two … like when we were doing the juggling, we didn't do every single step of that but I think we got from it what we needed to get from it. (Classroom teacher, intervention school)*

> *In year three where the learning of the new skill was British Sign Language, because I did that when I was at school, I could progress them on a bit further because I had an existing knowledge of that already, so I adapted it to sort of make it a bit more challenging for some pupils because the class I had last year were very able and they picked it up quite quickly. So, it was quite nice to be able to challenge them and push them on a bit further as well. (Classroom teacher, intervention school)*

The structured nature of the lesson plans, as described by teachers in the examples above, meant some teachers did not follow exactly point by point and/or adjusted the difficulty of lessons for some children.

The delivery of ReflectED and engagement with it (RQ 9a)

Findings summary

This section explores the delivery of different components of ReflectED and how participants have engaged with it. The data presented is drawn from teacher interviews, pupil focus groups, post-intervention surveys and structured observations. Synthesising evidence from these data sources illustrate a picture of delivery of and engagement with ReflectED where there were high levels of engagement, the performance tag colours emerged as particularly useful and well-engaged with aspects of the intervention (teachers described and reported that this was especially so for younger and lower ability children). The completion of reflections was described as more difficult, by some respondents in respect of its accessibility for lower ability pupils and those with additional needs. In terms of engagement, as previously discussed in relation to barriers, reflections were a less easy aspect of the ReflectED intervention to engage with. ReflectED lessons themselves, and the training provided by the development team (launch and hub meetings) alongside additional support were received positively by most schools and teachers, this would suggest that this assisted in the continued engagement of the intervention schools throughout the evaluation period.

All post-intervention survey respondents in intervention schools (headteachers, Year 2 and Year 6 teachers, and nominated leads; n=177) were asked their agreement with a series of Likert scale statements relating to their experience of implementing ReflectED:

- 95% (168 out of 177 respondents) agreed (67% strongly and 28% somewhat; n=118 and 50 respondents, respectively) that using ReflectED had been a positive experience, 5% (8 respondents) neither agreed nor disagreed and only 1% (1 respondent) somewhat agreed.
- 96% (169 out of 177 respondents) agreed (75% strongly and 25% somewhat; n=132 and 137, respectively) that ReflectED had increased their own awareness of metacognition.
- In addition, majority of headteachers (98% or 47 out of 48 headteachers) agreed strongly or somewhat, 77% and 21% or 37 and 10 respondents, respectively) that implementing ReflectED was a positive experience for their school. The remaining 2% (1 respondent) neither agreed nor disagreed.

Year 2 and Year 6 teachers were also asked some additional questions relating to pupil engagement (n=97).

- 94% (91 respondents) agreed (47% strongly and 46% somewhat agreed, n=46 and 45 respondents, respectively) that the pupils enjoyed ReflectED lessons; 4% (4 respondents) neither agreed nor disagreed; and only 2% (2 respondents) disagreed (1% somewhat and 1% strongly; 1 respondent each).
- 91% (88 respondents) agreed (45% strongly and 45% somewhat agreed; n=44 and 44 respondents, respectively) that ReflectED lessons had had a positive impact on pupil learning, and 9% (9 respondents) neither agreed nor disagreed.
- 76% (74 respondents) agreed (27% strongly and 50% somewhat; n=26 and 48, respectively) that completing reflections as part of ReflectED was positive for pupil learning, 20% (19 respondents) neither agreed nor disagreed and 4% (4 respondents) somewhat disagreed.

Post-intervention survey respondents in the intervention schools (n=177) were asked if they planned to continue using ReflectED in the following academic year: 78% indicated that they did plan to continue; 18% stated maybe; and only 5% said no. Of the eight respondents who indicated they were not planning on using ReflectED the following year, three respondents indicated that they would still use some of the approaches, and the remaining five respondents cited other pressures and commitments including staff turnover, an already full curriculum, and the decision of the schools governing body to not continue with the intervention.

Qualitative data from semi-structured teacher interviews and pupils focus groups provided insight into how both teachers and pupils have engaged with the intervention, specifically how pupils have engaged with: i) the ReflectED lessons; and ii) writing reflections and the use of tag colours.

Most teachers reported that most pupils liked and were highly engaged by the skill based ReflectED lessons and found them enjoyable. Indeed, the structured observation data suggested that children were largely engaged in the ReflectED lessons that were observed. Observers assigned a number (3 = often, 2 = sometimes, 1 = rarely, 0 = never/not seen) and of 16 observations 4 were assigned a value of 2 for 'Children are engaged' and the other 12 observations a value of 3. Such accounts were supported by comments from pupils during focus groups, where the ReflectED lessons were often described as being fun and generally not in need of any changes.

> *Mine have just really enjoyed the learning a new skill, and they're so proud of themselves for being able to tie their shoelace. And my son's in Year Two and he can't, that's like wow, but there's such a simple process. And they're children that, you know, you maybe sometimes have lower expectations than you should of them, and actually I'm really surprised with how many of them can actually do it now. It's not reinventing the wheel or anything but they can actually just do it, and I think it's because they've been so engaged with it. I don't know what it is but they're just really engaged with the sessions that we do and, you know, they're really focused when you're talking … I think it's because we're honest with them as well and we, like, we can't do it. And they really like those stories of things that we say that are absolutely true, and they can see that. That makes them feel secure and be able to talk about it as well. (Classroom teacher, intervention school)*

> *I've got a couple of behaviour issues, but what class hasn't? I do find in the actual ReflectED sessions, you know the weekly … they just, they're hooked because they're so engaged … (Classroom teacher, intervention school)*

> *I like my reflective lessons because it's fun, and I really like writing and drawing pictures. (Pupil, intervention school)*

Although, some teachers highlighted that some groups of children, e.g. those with SEND and/or of lower ability, were less able to engage with some of the skills based ReflectED lessons as they were too complex, sometimes leading to frustration among these pupils.

> *… the shoe laces … it was too hard for Year 1. I think for the rest of them…actually they've all been fine. You know, the chopsticks for Year 2 was perfectly pitched, the problem with the shoelaces was that there were children that have fine motor difficulties, they are already at a disadvantage? So, there's a couple of little boys in my class that can't do up buttons yet, so they haven't … they didn't meet the Early Learning goal of moving and handling to do up their buttons and things, the social care one, so it was hard for them, and I also already had about three kids that could do it. And they found it … they liked the idea of teaching other kids, but they were a bit bored with it. I mean it was so hard, I think it was one of the hardest things I've ever had to teach kids. We had tantrums, we had throwing shoelaces across the room. (Classroom teacher, intervention school).*

Several teachers explained that having a lesson dedicated to metacognition/ReflectED had generally helped to facilitate engagement with ReflectED, and metacognition more generally, at both a pupil level and teacher level.

> *I think it's the structured way in which the lessons are very clearly focused on the next bit of metacognition, if that makes sense? And it's focused then on developing those skills as opposed to developing a skill or teaching a concept or some knowledge that as a teacher you'd naturally teach, but it gives you … and I remember thinking, was it the third or fourth lesson … gosh, I'm doing it … you know, it seems a bit slow almost, but actually now I've been … by the fifth lesson, I was thinking, well, actually I understand this now, whereas maybe the first few lessons I was unsure because it seemed quite a laborious way of learning something, but actually it was because you're learning not just the skill of using a compass but you're also learning that skill of reflection or developing the skills the children already had. (Classroom teacher, intervention school)*

> *[the pupils] have been positive about it [ReflectED lessons] and I think because we have really gone to town, you know, the metacognition discrete lessons that we've got, it's enabled them just to have like an explicit lesson of just learning about that and then that's helped definitely to filter it out. So that's made them feel very knowledgeable and confident when they talk about it. Without that discrete lesson*

from the [inaudible 11:01] I don't think ... they wouldn't be as able with the other lessons. (Classroom teacher, intervention school)

Most teachers described how their classes actively engaged with the task of reflecting and writing reflections, and during the focus groups, some pupils discussed the benefit they felt that reflecting had on their learning.

I can see why it's helped me a lot because during when I've done reflections, I'm just like, I know when I could do that some, another time and then get it stuck in my brain on how to do it. (Pupil, intervention school)

Despite this, teachers acknowledged that reflecting did not come naturally to most pupils. Teachers explained that many children do not find reflecting easy and recognised that it was a skill that needed to be continually taught and developed. Teachers described that sometimes children's reflections were surface level.

I think for me reflection, having the kids, the kids being able to reflect on what they're doing, how they're doing it, why they're doing it, and it's not just a task. It's how I'm doing this task. And because we're so, we're further up the school they can really unpick what it is I had to do to get to this part. How am I going to move to my next stage, what is it I need to do? So, it's them being able to really reflect on their learning. (Classroom teacher, intervention school)

[Reflecting is] just part of their every day. Which is great. But yeah, that honesty is tough for some of them. (Classroom teacher, intervention school)

That's an area that I think we need to keep developing because I think children per se think…it's sometimes quite brief or it's, oh, I don't know, I don't know what I could do to … you know? So that is an area that I personally think I need to develop and become better at. (Classroom teacher, intervention school)

Teachers once again spoke about the challenges they had engaging their younger cohorts and those of lower ability to reflect using the *ReflectED* approach, which they felt was unrealistic and should have been simpler while introducing the children to the new approach.

I think especially with Year 1, I think we went in a little bit too soon with them reflecting on like a formal strip of paper, so we trialled what could you do before the lesson, what could you do after the lesson kind of thing. They had to write a short sentence for each. But in Year 1 at that point we introduced it that took a whole lesson to do it. It's just unrealistic I think. Again, we ReflectED and adjusted the coloured dots. That was far, far more appropriate for Year 1 and Year 2 at that point. (Classroom teacher, intervention school)

Our lower [ability] children have found it really hard haven't they, but the highers [higher ability children], they always say, because – so they give reasons as to why they are where they are. It's not just, I can't do it. It's, I've never done this before so, and it's new to me, so therefore ... and they do give a reason as to why. It's just the lowers [lower ability children] have found it quite tricky. (Classroom teacher, intervention school)

When you get them to do written reflections, I think that down here it's taken a lot longer to get them used to doing it and I would say that's probably one of the things we haven't done so much because it's harder, so we've been doing a lot more of the recorded reflections, but yes, we've got a lot of children in the class who aren't used to writing a lot because it is a lower group, like I've got two Year Three children in that class, one can't even write her own name yet, so to be able to get them to write a reflection, I've been finding that quite hard. So it is better when they do record it. (Classroom teacher, intervention school)

Teachers discussed how such children struggled to reflect as they did not possess the required language and/or understanding to do so, however the use of tag colours enabled them to engage with the concept of reflecting on their learning, which was viewed positively by teachers.

I think accessibility because some of my, kind of, lower abilities with not much language weren't really able to reflect. I think the colours really, really helps them do that, so I'd say for me that's a big positive. (Classroom teacher, intervention school)

Yeah, they can do. I often find with lower ability children, sometimes, they'll start on red, and they'll put themselves back on red at the end of the lesson. And I'm like, but … and then we talk about it, and in actual fact, they've achieved what they needed to do. Yeah, but I found it really hard. Yeah, you might have found it really hard, but you've actually done it!. (Classroom teacher, intervention school)

… those who don't have much language at all; they understand what the colours mean, and even though they aren't necessarily able to say what helped them but they are able to think about where they are at in that particular learning. (Classroom teacher, intervention school)

In the interviews, teachers discussed how using tag colours helped all children engage with reflecting, regardless of ability, and pupils described how they found it useful. This aligns with data collected via the structured observations of ReflectED—in 14 of the 16 lesson observations the evaluation team saw evidence of the performance tag colours in the classrooms where they observed lessons. Teachers highlighted the importance of embedding tag colours as a metacognitive tool within the classroom for children to engage with it fully to assist their learning.

I think what … I've heard … initially I heard round school that some children were really quick to jump in at blue, particularly the younger children, because I think they felt they couldn't be anything but blue. But I think the more work we've done now, the children are realising it's okay to be red. You're here for a reason, you're in school for a reason, and even … I mean I know I talk to my children about it, that even when you get to being an adult, you're not going to get everything right. It's not … life's not like that. ... Some of the children who would have … I would have said given up really, really easy, the tags have actually helped them to realise, actually there's some things I am a green on. It surprises them actually, because they might be the children who, beforehand, would have always gone, I can't do anything. (Classroom teacher, intervention school)

I think mine now have a really clear understanding and an appreciate, actually getting to blue isn't what it's about. And they all know that when I sit down on an evening and open their books not everybody will have, kind of, progressed to blue, and it's actually the journey that's important. So, some of them are red and they know they're on red so they go home and practice big maths, more big maths, more times tables, come back and then try and say, get themselves to yellow. They'll say to me, I think I've got to yellow now because I've just got a little bit more understanding of it. And so, I think there's a good awareness of the colours and they've responded really well. (Classroom teacher, intervention school)

Reach and perceived or actual benefits (RQ11)

Findings summary:

Teachers perceived benefits and reach of the ReflectED intervention for a range of pupils, we note that there was often particular reference to the utility of performance tag colours to facilitate reflection and the 'visibility' of learning/progress where pupils can see the 'bigger picture' of their own learning. The value of seeing and learning from mistakes is also noted, with data from teacher interviews and pupil focus groups describing ReflectED as beneficial for resilience towards learning. In terms of specific groups of pupils benefiting, almost half (49%) of the respondents to the post-intervention survey did not believe that ReflectED was beneficial for a particular subset of pupils, the 51% who did attribute the benefits they perceived for pupils of a particular level of academic ability—this was not consistent across the responses with 33% describing more perceived benefits for higher ability students, 27.8% for middle ability students, and 23.7% for lower ability students. This finding is perhaps not surprising given the focus in ReflectED on pupils working in mixed ability pairs.

During interviews, teachers were asked if they perceived there to be any benefits to pupils by delivering and implementing ReflectED within their classrooms/schools. Many teachers discussed how they had noticed changes in most children's behaviour towards learning and the language they used in relation to metacognition. Teachers explained that over time ReflectED had equipped children with a new set of tools (with particular reference to the usefulness of tag colours) and metacognitive language to approach and engage with new and/or complex tasks. As a result of this, teachers observed improvement to children's communication and a boost to their self-esteem and resilience towards

learning. Teachers reported that pupils had changed their mindset in relation to making mistakes and were more accepting of them and more encouraged to overcome them as a result of ReflectED. There were some examples of how teachers and pupils had altered the terminology of a 'mistake', and since referred to them as 'amazing mistakes', 'marvellous mistakes' or a 'good failure'.

> *They're [the pupils] more able to now see that they can't just not do something big picture, they can't do one strand of it. I think that's a thing that's massively boosted their confidence and self-esteem and the willingness to carry on because they're not feeling, 'I've failed everything', they're just thinking 'it's just this bit' and that's smaller for them to deal with. I think they're more resilient now with tasks that would previously perhaps have set them up to fail, so if it's going to be something quite hard in the lesson they're more willing now just to keep going rather than just to have kind of maybe behavioural breakdown or crying or acting out. There is a lot more ... you hear them talking about, have another go, I can't do it yet. So, to me that's them thinking, I'm keeping going I can do it. (Classroom teacher, intervention school)*

> *They know the language of perseverance, resilience; they know what trial and error is. They know good failure. All these words, they say independently without prompting. That's because of discrete lessons and then also obviously modelling that in normal lessons as well. (Classroom teacher, intervention school)*

> *But I do think it's had an impact, especially with like mistake culture if you like. A lot of the children, especially the higher ability children in year six were really struggling when they made a mistake and really punishing themselves for it. But now we go with like ... I'll have a look round and I'll like find a mistake and I'll get them to come to the front, like 'right this is an amazing mistake' and they'll share their amazing mistake and what they've done and we'll put it on the working wall and they love it. But whereas before it was like, they'd be gutted to make a mistake, whereas now they're quite happy to share it. (Classroom teacher, intervention school)*

Similar perceived benefits were corroborated by pupils themselves within focus groups. Pupils explained how their own attitudes towards making mistakes within and outside of the classroom had changed because of ReflectED. Pupils described themselves as being more resilient, more likely to persevere with tasks and view mistakes as a positive part of learning as opposed to a negative.

> *I feel that ReflectED has helped me become more resilient because before we did it, if we were to do a trial and error work, I'd do five, and then if that wasn't right, I'd kind of sometimes give up, like I said I didn't understand it anymore, but now that we've done ReflectED, I've learned to be more resilient and things. (Pupil, intervention school)*

> *... in ReflectED we learn like all our mistakes and that it's good and we can put it with all our other lessons. (Pupil, intervention school)*

Teachers also observed that pupils were now more inclined to ask them for help, which teachers perceived to be a result of ReflectED. This was confirmed within pupil focus groups, as pupils discussed that the ReflectED approach made it easier for them to communicate with their teachers—often through the use of 'tag colours'—when they needed help with something, which the pupils viewed as being beneficial.

> *Yeah, asking for help, wanting extra help, asking for extra homework, discussing with their partners what do I need to do, where am I going wrong. So, using their peers as well as the adults in the class. (Classroom teacher, intervention school)*

> *I think ReflectED is good for us because teachers can see what we're thinking inside our head and it's like talking to them but in writing. So like, they can see that, oh this person needs a bit more support on this so like, if they're blue, they're red, yellow. If they are green they are confident and if they're yellow they might need a bit more help and if they're red, they don't understand. And I think it's really good because it shows what we're thinking and how we think and so they can see it. (Pupil, intervention school)*

> *I think that using different colours is a great help because if you don't understand something, like if we're doing something hard and something you've never done before you just put down … when I started this lesson I was blue, but if you write that down, the teacher will think that you understand fully and you can do more challenges but you actually don't know what to do so the colours they help you to take different steps in learning what to do. (Pupil, intervention school)*

Although teachers were observing differences in children's language and behaviour towards learning, one school noted that these changes did not translate into improvement in pupils' academic outcomes at a class/school level. In the interview extract below, the teacher discussed how it was difficult for the school, as a whole, to quantify the benefit of ReflectED.

> *I think what we found hard is the fact that showing the progress in terms of numbers at this stage has been really difficult. We have seen the improvement in the language the children are using and a lot more of those kind of not data driven ways, and I think it will take time to see it on data across time to be able to actually look at the impact and how much progress those children have made, and I think from like, I know it's difficult in terms of from my position having to report to the Head, because I think she literally wants me to show all this massive improvement and it's suddenly going to like change the world and actually I think it's more about like you say, the fact that ... Their attitude to learning, the language that they are using, the kind of feel that you have in school that mistakes are okay as long as it's part of that learning process, and I think that's the definite shift that we have seen more than, you know, Joe Bloggs is … (Classroom teacher, intervention school)*

One teacher considered that ReflectED encouraged children to self-reflect and had improved behaviour within their classroom.

> *I would definitely say maturity levels, behaviour levels. They weren't poorly behaved, but what I have noticed is there's a lot less typical fallings out, because so and so has said this, or so and so … so and so has done this and I don't … I think that … the level of maturity and the way that they're reflecting on their own. I'm not just saying their learning, but themselves, I think, has really helped them. I'm not saying it's perfect all the time, because what classroom is, but I do think actually…because we talk about it across the range of things and I give them so many examples. Like the lesson that you'll see, I've actually geared it around last night's match. So, I'm hoping that you'll see it in terms of, you know, something that they're interested in. But I do … I personally think, looking at my children, maturity and reflecting I think … it's been the reflecting side, more than anything else… (Classroom teacher, intervention school)*

During semi-structured interviews, teachers also considered the impact the perceived ReflectED had had on specific groups of pupils (e.g. those with SEND and EAL) although the feedback was mixed. Some teachers reported that these groups of children gained the greatest benefits from the school implementing ReflectED, whereas other teachers observed lower engagement levels among these groups (we present the data in relation to the latter in research question 9a).

> *I think I've seen the biggest on my SEND children, and my children who had, their progress had stalled or they were making very, very slow progress. I think one) that's because of self-esteem because I think when they get to blue they think they've moved mountains, and that's really quite lovely to see, and they, it just empowers and then they progress more. So, I think that whole part of self-esteem is really quite amazing. And then I think secondly, it breaks it down and it makes it really clear for a child to be able to explain their learning, and I think a lot of children struggle to explain themselves. I think it's improving their communication. (Classroom teacher, intervention school)*

Year 2 and Year 6 class teachers in intervention schools were asked in the post-intervention survey if they felt that some groups of students had benefited more from ReflectED than others. The results are presented in Table 21. Of the 97 respondents to this question: 50% (48 respondents) indicated that they felt the intervention did not benefit any particular group of pupils; 12% (12 respondents) indicated they felt boys benefited; 12% (12 respondents) also indicated they felt girls benefited; 33% (32 respondents) indicated that higher ability pupils benefited; 28% (27 respondents) indicated middle ability pupils benefited; 24% (23 respondents) indicated lower ability pupils benefited; 10% (10 respondents) indicated pupils with additional needs benefited; and no respondents indicated EAL pupils learning English as an

Additional Language benefited. With regard to the 'performance tag colours' and reflection, teachers frequently spoke (in interviews) about confident pupils, particularly at the beginning of the intervention period, reporting that they could do everything (i.e. that they were green or blue), and less confident pupils would report red. Teachers described that over time, during the intervention period, pupils learned how to map the colours to their learning trajectories. For example:

> *… "you noticed, you know, they [the children] went to the red…I was like gosh, you're not red. And it was just because well I still can't do it, so I'm frustrated about it. But they know how to do it, and actually a lot of them would really, for me, if I watched them do it, they are confident with it, but their…their sense of self tells them … yeah, there are a couple with confidence issues." (Classroom teacher, intervention school)*

> *"I heard round school that some children were really quick to jump in at blue, particularly the younger children, because I think they felt they couldn't be anything but blue. But I think the more work we've done now, the children are realising it's okay to be red. You're here for a reason, you're in school for a reason, and even … I mean I know I talk to my children about it, that even when you get to being an adult, you're not going to get everything right. It's not … life's not like that ... Some of the children who would have … I would have said given up really, really easy, the tags have actually helped them to realise, actually there's some things I am a green on. It surprises them actually, because they might be the children who, beforehand, would have always gone, I can't do anything. You know, I can't … and I think the initial activity we did, we tried what Rosendale showed us to do, where they put the tags up and some … I think a lot of the classes tried that, didn't they?" (Classroom teacher, intervention school)*

Table 19: Groups of students that teachers reported as particularly benefiting from the ReflectED intervention (n=97)

Pupil group (as listed on post-intervention survey)	Number* (%)
Girls	12 (12.4)
Boys	12 (12.4)
Higher ability students	32 (33.0)
Middle ability students	27 (27.8)
Lower ability students	23 (23.7)
Students with additional needs	10 (10.3)
Students who have English as an Additional Language	0 (0.0)
No groups of students in particular	48 (49.5)
Other: • Free-text responses: • Lower ability students have gained confidence in being able to recognise areas of the curriculum they have strengths in • Pupil Premium learners • I think some of the lower ability learners struggled to reflect honestly as it was easier to repeat what they always write in their reflections. Their oral reflections with an adult questioning were always the most effective • I think they have all benefited and been able to take something from ReflectED lessons and apply it to their learning • I think everyone has benefited equally	5 (5.2)

* More than one response could be given.

A total of 49 free-text answers were provided by classroom teachers describing why they thought specific groups of students saw more benefits from ReflectED, they were analysed inductively. Around 14 responses included comment(s) about the perceived impact on pupil confidence, resilience, or independence of ReflectED, this aligns with the teacher interview data where analysis identified a teacher-reported increase in pupil self-esteem and resilience towards learning. The first example below is particularly salient, it also includes reference to the teacher identified benefits of the

'performance tag colours'. The second example illustrates another common theme in the inductive analysis of these free-text responses and the teacher interview data—a perceived increase in pupil ability to see (and reflect on) their process(es) of learning.

> *Working in mixed attainment groups has enabled these children to work alongside children who can coach and support them with their learning. They have been exposed to so much vocabulary and using the stem sentences has really improved their oracy. The low attaining children have began to use strategies to help them to be successful and have developed the independence to access resources without the need of an adult. For the high ability children, they have benefited from explaining their understanding of a concept, in order to coach others. These children are also seeking challenge independently. (Classroom teacher, intervention school)*

> *Lower ability children have been able to improve their skills of verbalising which colour they feel they match to and why. However, the middle and higher ability children have seen more benefits as their vocabulary, understanding and Independence of self reflection has improved greatly. They are able to support lower ability children too. (Classroom teacher, intervention school)*

In relation to explaining why specific groups of students have seen more benefits from ReflectED teachers also commented on the language associated with the ReflectED intervention and the perceived benefits for pupils who teachers felt were more able to *'access the vocabulary' (classroom teacher, intervention school)* that allowed them to then talk about their learning, with both teachers and peers.

Evidence of metacognition and change across the intervention period (research question 13)

The data that was originally planned to be presented in relation to research question 13 (the pupil reflections) was, as previously explained, superseded by self-reported data from the post-intervention survey for intervention schools and is supported by data from teacher interviews and pupil focus groups in this section.

Data sources:

- self-reported post-intervention teacher survey;
- teacher interviews; and
- pupil focus groups.

Findings summary:

Data collected that focused on exploring for evidence of metacognition does suggest that teachers believed that there was some evidence of an improvement in the quality of pupil reflections over the course of the intervention period. We note that the expected number of reflections (as per logic model) was not met across the intervention schools. When asked what evidence of metacognition looked like, teacher responses (and pupils supported to some extent—this was less the focus of pupil focus groups and more teacher interviews) focused on use of metacognitive language and how pupils were talking about their learning after input in relation to the ReflectED intervention. As previously mentioned, in relation to engagement, there was again a particular focus on the usefulness of the performance tag colours and Growth Mindset focused teaching.

All respondents in intervention schools to the post-intervention survey (headteachers, Year 2 and Year 6 teachers, and nominated leads; n=177) were asked their agreement with a series of Likert scale statements relating to pupil's talk about their learning and their metacognition.

- 98% (174 respondents) agreed (70% strongly and 28% somewhat agreed; 124 and 50, respectively) that ReflectED supported pupils to talk about their learning, and 2% neither agreed nor disagreed (3 respondents).
- 97% (172 out of 177 respondents) agreed (72% strongly and 25% somewhat; 127 and 45, respectively) that ReflectED supported pupils to be metacognitive, and 3% (5 respondents) neither agreed nor disagreed.

Year 2 and Year 6 teachers in the intervention schools were also asked about perceived evidence of metacognition among their pupils, particularly in relation to the reflections that pupils completed as part of the ReflectED intervention.

- 40% (39 out of 97 respondents) definitely felt that they saw evidence of metacognition in pupil reflections about their learning and a further 45% (44 respondents) felt that probably they saw such evidence, 14% (14 respondents) indicated that they might, or they might not.

- Of the 83 respondents who felt they either did or probably saw evidence of metacognition in pupils' reflections, 80% (66 respondents) saw this evidence in pupils written reflections (either on paper or in Seesaw) and 80% (66 respondents) in pupils' verbal reflections. One respondent indicated that this evidence had been borne out in pupil voice interviews.

- In addition, 53% of Year 2 and Year 6 teachers (52 out of 97 respondents) strongly agreed with the statement 'Pupils can confidently assess their learning using the performance tag colours', 41% (40 respondents) somewhat agreed 4% (4 respondents) neither agreed nor disagreed and 1% (1 respondent) somewhat disagreed; and

- 33% (32 respondents) strongly agreed with the statement 'The quality of pupil's reflections has improved over time', 57% (55 respondents) somewhat agreed, 9% (9 respondents) neither agreed nor disagreed and 1% (1 respondent) somewhat disagreed.

The post-intervention survey asked teachers in intervention schools what the evidence of metacognition looked like (with a prompt of 'For example, do they use specific language or talk about particular things') and 83 free-text responses were inductively coded. There was a focus on talk and language, both about and for learning in the coded responses. A total of 50 (out of 83) responses in total attributed evidence of metacognition to their talk and/or language about learning. Respondents described pupils as being more able to or more accomplished at talking about their learning, and in their use of language about learning (e.g. metacognitive language, language, and phrases specific to the ReflectED intervention and language in relation to target setting). For example:

> *I have both seen and heard my children talking to each other without prompts and using the vocabulary. (Classroom teacher, intervention school)*

> *In their use of vocabulary when talking about their learning and in reflections. (Classroom teacher, intervention school)*

> *Some children use specific language relating to the colours… (Classroom teacher, intervention school)*

The colours mentioned in the last quote refer to the 'performance tag colours' that are part of the ReflectED approach, where pupils are facilitated to consider and reflect on where in their learning they are. Of the 83 respondents, 22 respondents described evidence of metacognition as being present in pupil use of the tag colours. In relation to the tag colours, teachers described these as being helpful for pupils in articulating where they were in their learning. For example:

> *Talking about colours and why they think they are for example 'blue'… (Classroom teacher, intervention school)*

> *The children were able to describe their understanding of a task more accurately, instead of claiming to be experts (blue) at things they had never done before! They could discuss a true starting point and what the next steps to success were. (Classroom teacher, intervention school)*

The latter example, links to another frequently described type of 'evidence of metacognition' that teachers described in these free-text responses—pupil awareness of and reflection on their process of learning. The process of learning content in the free-text responses given also link clearly to talk and language about learning. For example:

> *Children have found it easier to articulate their learning journey and have been able to provide examples and explanations. (Classroom teacher, intervention school)*

Alongside the metaphor of the learning journey, the metaphor of the 'learning pit' that is included in the ReflectED training materials was also cited as a means of evidence of pupil metacognition as linked to the process of learning, e.g. 'Children are talking about the learning pit, they use the language of I'm not there YET and they understand that specific feedback is helpful'. Many of the responses included two or more of these aspects of 'evidence of metacognition' (e.g. process and talk/language about learning), the 'YET' in the latter example refers to Growth Mindset (Dweck, 2016). Another example illustrates how often teachers reported a mixture of different 'evidence of metacognition' post-intervention, in this example Growth Mindset sits alongside tag colours and talking about learning:

> *They use specific language such as growth mindset vs fixed mindset and are able to tell why they are a certain mindset. Using the tag colours as sentence starters and give reasons why they are that colour. (Classroom teacher, intervention school)*

In line with the data analysed above from the post-intervention survey, during semi-structured interviews teachers reported that they had observed an increase in pupils' understanding and use of metacognitive language over the course of delivering ReflectED.

> *They know the language of perseverance, resilience; they know what trial and error is. They know good failure. All these words, they say independently without prompting. That's because of discrete lessons and then also obviously modelling that in normal lessons as well. (Classroom teacher, intervention school)*

Teachers explained that they perceived this to be a result of the implementation of tools (such as 'tag colours'), and concepts (such as 'growth mindset') that underpin a metacognitive approach to learning. Together, teachers reported that these assisted pupils to communicate more effectively about their confidence and ability in approaching and completing a task by helping them breakdown and identify parts of the task they are struggling with, ultimately supporting them with their learning.

> *I can see an impact even with my class, and they are quite a low ability class. They never would have been able to talk about their learning before, so it has had a real positive impact in my class. (Classroom teacher, intervention school)*

> *I've taught Year Six for a number of years, and I think that they now have the language to use and the tools in a way and whereas if children were stuck before, they might … you know, I find children to be very upset about that and they don't exhibit any of that and actually, you know, and they're quite happy to say I'm on red, I don't understand, but they're more specific in what they don't understand than they would have been historically. (Classroom teacher, intervention school)*

> *Since we did the Growth Mindset unit, that's made a huge difference because I've noticed they, kind of, throw the words around, you hear the words going around. And, I mean, even … You could hear them when we were in the class just then, like using words that they're not necessarily going to use. It's made them think as well because they've had assemblies on Growth Mindsets, so it's kind of being reinforced. And then if one child hears someone saying something and I'm, oh well done, then somebody else wants to get that well done; so they kind of all start using, and using it in the right context as well. (Classroom teacher, intervention school)*

Teachers also discussed that pupils were exhibiting metacognitive skills learned within ReflectED to other subjects/lessons. Although teachers recognised that children may not always be aware that they are doing so, they acknowledged that older children were more likely to be conscious of doing so in other subjects to progress their learning, in comparison to younger children.

> *… the kids further up the school especially are quite confident in, like, applying it [the ReflectED approach] to different things maybe and talking about it.(Classroom teacher, intervention school)*

> *They [pupils] use the skills. I don't know that I would necessarily … if they would identify, if we were in a maths lesson, say, I don't know that they would identify, oh, I'm using metacognition here … so explicitly. But I think they do use the tools from it; because we talk about building a toolkit all the time, and taking it to other lessons. That's, sort of, the analogy that we use. And I think they do use them, but*

I don't know that they would know they were using them, if that makes sense. (Classroom teacher, intervention school)

The children didn't necessarily see it as metacognition, but actually, what they're seeing now is, they're transferring those skills that we're doing in ReflectED metacognition, through that. But I do think … like some of the boys, in particular, are thinking about what they do outside of school and they talk a lot more about that as well. (Classroom teacher, intervention school)

I think some better than others but yes, we do see it coming through other subjects. And we do, as teachers, encourage it as well if it's not there but it's needed. (Classroom teacher, intervention school)

They use the skills [taught via ReflectED in other lessons]. I don't know that I would necessarily … if they would identify, if we were in a maths lesson, say, I don't know that they would identify, oh, I'm using metacognition here … but I think they do use the tools from it; because we talk about building a toolkit all the time, and taking it to other lessons. That's, sort of, the analogy that we use. And I think they do use them, but I don't know that they would know they were using them, if that makes sense. (Classroom teacher, intervention school)

Data from pupils focus groups supported the observations of teachers, as pupils described using the skills developed via ReflectED in other subjects and acknowledged the usefulness of doing so to support their own learning. Pupils talked about how they were able to transfer lessons they had learned in ReflectED lessons across different curriculum subjects including English and mathematics and how metacognition had been helpful in terms of challenging 'fixed mindset'. There is a sense of pupils' metacognitive journeys in this data, across the period of the intervention.

Cost Evaluation

For the purposes of the trial, schools received subsidies from the EEF to cover various aspects of cost. Here we report costs associated with implementing the ReflectED programme in a 'real-world' context.

We estimated the average cost per pupil per year for schools implementing the programme. Although completion of the report and subsequent publication has been delayed due to Covid-19, this report was commissioned in 2017 and the evaluation was designed with the EEF current guidance in mind. Consequently, we have used the 2016 cost guidance (EEF, 2016), while accounting for the updated guidance issued in 2019 (EEF, 2019) where possible. ReflectED is intended to be embedded as a whole-school approach hence schools were asked to deliver the programmes to all year groups. The cost evaluation has been prepared under the assumption that the programme would continue to be delivered in this way. We have costed from the perspective of the average school included in this effectiveness trial (average school size n=318). Most of the data used in the cost evaluation was drawn from the post-implementation survey of headteachers in the intervention arm (n=48). Other costs were informed by our knowledge of the programme.

Table 22 details the resources needed to implement the programme. The primary sources of cost for implementing the programme are:

- training;
- staff time;
- licence fees; and
- materials.

Table 20: List of resources

Category	Item
Personnel for training	Implementation team
	School headteachers and a lead practitioner per school
	School staff (typically teachers)
Personnel for preparation and delivery	School staff (typically teachers)
Programme costs	Programme licence fee
Facilities, equipment, and materials	Information Technology facilities: tablets and internet connection. Printed resources
Optional materials	Schools could select additional materials to support learning such as juggling balls or other lesson props

To deliver the programme, there are some pre-requisites in terms of Information Technology (IT) facilities. These could constitute start-up costs if schools decided to purchase these specifically to deliver the programme, however, the IT requirements were stated on the memorandum of understanding for the purposes of this trial. Despite this, 10 schools in the intervention arm (18%) reported purchasing tablets specifically to deliver ReflectED. It is assumed all other schools used existing IT equipment.

Time

Prior to implementation, staff time was required for training. The headteacher as well as the nominated 'lead practitioner' from each school were initially required to attend a 1-day training event. Throughout the implementation period lead practitioners were also invited to attend a 1-day hub meeting once per term (five school terms in total over the course of the trial). Although it was envisioned that the practices of the ReflectED programme would be embedded into teaching methods, schools were asked to deliver formal ReflectED sessions with each class once per week, at around 20–40 minutes per session for a whole academic year of 39 weeks, amounting to 13–26 hours of intervention time in total. In the end of trial survey, schools were asked to what degree they agreed with the statement 'ReflectED lessons can be taught in the suggested 20–40 minutes'. Only 13 respondents out of 100 (13%) disagreed with this statement. Of those disagreeing, the reported mean time spent delivering sessions was 48 minutes. To be conservative, we have calculated time costs based on the upper end of the pre-specified time estimate, hence, sessions are assumed to take 40 minutes on average (Table 23). Sessions ran during a typical school day during typical lesson time.

We did not capture any data regarding additional time needed outside session time to deliver ReflectED. No other time costs were involved.

Table 21: Total time devoted by school staff to deliver the programme over 39 weeks per year

Activity	Year 1 Mean number of hours	Year 2 Mean number of hours	Year 3 Mean number of hours
Initial training	6.5	0	0
Hub meetings	19.5	19.5	19.5
Session delivery per year group	26*	26*	26*
Session delivery per school (Reception year to Year 6)	182*	182*	182*
Initial training	6.5	0	0

Note: * Assumes a mean session delivery time of 40 minutes, once per week for 39 weeks for one class per year group. No information on additional time costs was captured.

Financial costs

Training

For the purposes of the trial, a 'cascade' model was adopted whereby the implementation team trained a lead practitioner as well as the headteacher within each school who in turn trained teachers. A similar model could be used to roll out the programme nationally.

Training of the lead practitioners/headteachers was delivered by staff from the implementation team at Rosendale Primary School (Lambeth, UK) through a 1-day event. We assume that the Rosendale Primary School team cost £185 per day. Lead practitioners within schools were not paid any additional fee to support training within school.

Lead practitioners/headteachers were trained in groups. A total of five 1-day training events took place between January 2018 and February 2018.

In addition to the training events, lead practitioners were expected to attend hub sessions once per term. 'Hub schools' were paid £1,000 per hub event to host these sessions. Throughout the implementation period intervention schools were provided with funding to provide cover for the lead practitioner to the value of £1,110 (£185 x 6 days release).

To our knowledge, there were no venue hire costs associated with training events or the hub meetings as these were held within schools.

Schools were required to pay their own transport costs to attend hub meetings. Data captured through the post-implementation survey of headteachers indicated that the average cost of transport amounted to £288 for the period of the trial, though this ranged from £0 (if staff were employed by a host school) to £900. We use the average reported cost for our costing model.

Materials

With regards to materials, to facilitate implementation of ReflectED, schools needed to have access to tablets and the internet. It was specified in the MOU that schools should already have these facilities, however, data from the post-implementation survey indicated that some schools did purchase additional tablets during implementation.

Schools delivering ReflectED were expected to print resources such as lesson plans and activity sheets. Data from the post-implementation survey indicated that schools spent on average £119 on printing (ranging from £0 to £500). We use the average reported cost for our costing model. As part of the post-implementation survey, headteachers (n=48) were asked about any other additional, or hidden costs, needed to implement the programme. Of the headteachers surveyed, 36 reported additional costs (75%). The mean additional cost was £1,787 though specified costs ranged from £0 to £25,750. Of those who reported the items paid for: 48% reported books; 60% reported practical resources; and 21% reported purchasing tablets. We use the average reported cost for our costing model (Table 24).

Table 22: Total costs of implementing ReflectED to 56 schools for the trial duration (18 months)

Cost type	Item	Start-up or recurring	Quantity required	Price per unit	Total cost
Personnel (training of headteachers and lead practitioners)	Trainer (implementation team)	Start-up	3 x staff for 5 days	£185 per person, per day	(£185 x 3) x 5 = **£2,775**
Personnel (hub events)	Trainer (implementation team)	Recurring	2 x staff for 33 hub events	£185 per person, per day	(£185 x 2) x 33 = **£12,210**
Personnel (staff cover)	Lead practitioner cover	Recurring	6 days cover per school	£185	£1,110 x 56 = **£62,160**
Programme costs	Hub host school costs	Recurring	10 schools	£1,000	£1,000 x 10 = **£10,000**
	Transport costs for hub attendance	Recurring	Per school	£288	£288 x 56 = **£16,128**
	Seesaw licence	Recurring	1 per school	£500	£500 x 56 = **£28,000**
	Additional materials (optional)	Recurring	Variable	Mean £1787	£1,787 x 56 = **£100,072**
	Printing	Recurring	Variable	Mean £119	£119 x 56 = **£6664**

Note: As most schools did not report purchasing Information Technology equipment, we have not included this as a start-up cost. Instead, this is factored into the 'Additional materials'.

School Level Costs

In total, £2,775 was spent on training the relevant personnel within the trial (headteachers and lead practitioners), which equates to £49.55 per school (total spent divided by the number of schools randomised). Additional costs included: the price of cover for lead practitioners, £1,110 per school; the price of the hub events including transport costs, £684.61 per school (calculated as the total cost of hub host school costs, the cost of the trainers to facilitate the hub meetings, and the cost of hub-related transport divided by the number of schools served [n=56]); the price of the Seesaw licence, £500 per school; and additional materials including printing, £1,906 per school.

Table 25 presents the total cost per year over 3 years. This calculation assumes an average school size of 318 pupils (as per our trial numbers), and that no additional training of school staff is required from the implementation team than as specified above. As some costs are dependent on the number of pupils, costs will vary somewhat by school size though this is unlikely to be substantive. The cost of the Seesaw licence has been costed based on the £500 given to each school during the evaluation to cover the licence fee, however, we are aware that this did not cover the full cost for many schools. The price of the Seesaw licence may vary somewhat under 'real-world' conditions where the fees are not fixed, and schools are required to request individual quotations. We assume £500 per annum per school for the purposes of this costing exercise.

Hub host schools were allocated £1,000 to cover the cost of hosting the termly hubs during the trial. It is assumed that each hub costs £200. All costs were captured for the whole trial period of five school terms, which crossed two school years. To estimate annual costs, total costs were divided by five to calculate a per term cost and then multiplied by three to calculate an annual cost (Table 26).

Table 23: Costs of the implementation of ReflectED per school per year for 3 years

Cost type	Item	Year 1	Year 2	Year 3	Total cost
Personnel (training of headteachers and lead practitioners)	Trainer (implementation team)	£49.55	–	–	£49.55
Personnel (hub events)	Trainer (implementation team)	£130.82	£130.82	£130.82	£392.46
Personnel (staff cover)	Lead practitioner cover*	£740	£555	£555	£1850
Programme costs	Hub host school costs	£10.71	£10.71	£10.71	£32.13
	Transport costs for hub attendance	£172.80	£172.80	£172.80	£518.40
	Seesaw licence	£500	£500	£500	£1,500
	Additional materials (optional)	£1,072.20	£1,072.20	£1,072.20	£3,216.60
	Printing	£71.40	£71.40	£71.40	£214.20
Total costs		£2,747.48	£2,512.93	£2,512.93	£7,773.34

* Assumes 4 days in Year 1 and 3 days (one per term) in subsequent years.

Table 24: Pupil costs over 3 years based on a school serving 318 pupils (assuming uptake of optional materials)

Number of years using ReflectED	Cost per pupil (£)
Year 1	£8.64
Year 2	£7.90
Year 3	£7.90
Overall	**£8.13**

Conclusions

Table 27: Key conclusions

Key conclusions
Pupils who participated in ReflectED made the equivalent of 1 month's less progress in the primary outcome related to Key Stage (KS) 2 mathematics, on average, compared to pupils in other schools. This result has a very high security rating.
Pupils who participated in ReflectED made the equivalent of no months' progress in the primary outcome related to KS2 reading, on average, compared to pupils in other schools. This result has a very high security rating.
Pupils eligible for Free School Meals (FSM) who took part in the intervention, made the equivalent of 1 months' progress in KS2 mathematics, on average, compared to pupils in other schools. They made no months' progress in KS2 reading, on average, compared to pupils in other schools.
Teacher surveys indicate that training led by Rosendale Primary School and cascaded in ReflectED schools was well received. Staff and pupils spoke positively about the programme, and they felt that it facilitated increased pupil awareness of their learning. Pupils in more than half of ReflectED schools did not complete two weekly online reflections. Verbal and written reflections were completed in addition to use of the online platform.
Post-intervention surveys from ReflectED schools suggested evidence of metacognition in pupil reflections. It was not possible to obtain surveys from all intervention schools, therefore this finding should be treated with caution.

We undertook a large and rigorous cluster RCT of ReflectED. There was little evidence of an impact on primary and secondary outcome measures. Indeed, the primary outcome was negative in that the control group made slightly more progress than the intervention group. There was a statistically significant interaction between the FSM pupil's mathematics outcome and the overall treatment effect. This was a 'qualitative' interaction in that for FSM children there was a slight, albeit not statistically significant, improvement in mathematics scores for the intervention children compared with the control children, in the context of a negative overall effect size. Because of the small effect size then these results are consistent with the hypothesis that it is unlikely that there is an educationally important difference between the intervention and control groups in all primary and secondary outcomes and for the FSM subgroup. Compliance with the ReflectED intervention was based on teacher attendance at training sessions and pupils completing two reflections a week, on average. While training sessions were relatively well attended (63% of schools sent one or more teachers to at least four of the five training sessions), engagement with reflections was lower than hoped. Less than half the schools reported that their KS2 pupils completed at least two reflections a week. However, the interpretation of this result is complicated by the fact that these data were obtained from the teacher post-intervention survey and some schools either did not respond or the survey was not completed by a KS2 teacher in the school. We were unable to obtain reflections data from Seesaw. Overall, only 14% of intervention schools fulfilled the compliance criteria, which inhibited a meaningful CACE analysis from being conducted. This analysis would have helped us to understand how effective the intervention might have been in schools with high engagement with the intervention.

Evidence to support the logic model

The results of the whole evaluation support some elements of the logic model, but other elements are not well supported. In summary the training and process elements were supported from limited to strong evidence obtained in the IPE. However, the impact elements were not well supported by the findings from the impact evaluation. Below, the items in the logic model are shown in italics, followed by a discussion of the strength of the evidence base from both the impact evaluation and the IPE, for support or not, of each item.

Training and cascading training; developing, and monitoring action plan; working with a partner; fidelity of implementation

There was strong evidence of a high level of attendance at the launch training across the intervention schools. Compliance to the intended model of training and cascading of training for the ReflectED intervention was perceived to be high. The training provided to headteachers and lead practitioners was very well received and enabled them to cascade it to the remaining teachers in their schools who also reported that it was well received. There is evidence that the regional hubs were well received and that they provided a good level of support (both peer support and support from the developers), sharing of good practice and appropriate input from the ReflectED team as required. The IPE provided quite strong evidence that ReflectED was being delivered as the intended whole-school intervention, both in terms of delivery across age groups and associated classes and within and between different subjects in the curriculum (i.e. not limited only to the weekly ReflectED lesson). In terms of fidelity of implementation in relation to the use of Seesaw to

complete reflections, it is unfortunate that despite technology requirements being part of the MOU that participating schools signed, not all had enough access (e.g. enough iPads or similar, and reliable enough internet connections) to consistently use Seesaw to complete reflections. The IPE data collected post-intervention also suggested that there was a preference in some instances for using paper-based reflections and the view that this was easier and more accessible.

Process: Increase in metacognitive approaches, discussion and reflection, pupil confidence, independence, and reflection (self-assessment); decrease in pupil disruptive behaviour

There is limited evidence of increase in metacognitive reflections in the pupils as reported by some teachers. The limitations in evidence here are partly due to the lower number of reflections being completed weekly by pupils than was expected. Data collected (in the post-intervention teacher survey) suggested that there was some evidence to suggest teachers believed that there was an increase in the quality of pupil reflections over the course of the intervention; this evidence is teacher-reported, and it was not possible for the evaluators to assess pupil reflections for evidence of metacognition and change in this.

There is quite good evidence that teachers observed improvement in children's communication skills, self-esteem, and resilience towards learning. There was good evidence in the IPE of teacher (and pupil) reported improvements in relation to use of language about learning and pupil ability to articulate and talk about their learning process—this links to the aspects of the logic model that focus on 'talk around metacognition + knowledge and understanding of good learning'.

Pupil outcomes

There is strong evidence from the impact evaluation that there were small negative effects of 1 month of less progress of the ReflectED intervention on: Mathematics KS2 primary outcome and no effects of zero months' additional progress on Reading KS2 primary outcome. Attrition was low and slightly differential (slightly greater in the control group) but multiple imputations did not materially change the results. The secondary outcomes demonstrated that there were no effects of zero months' progress for the JrMAI, Mathematics KS1, and Reading KS1, with small positive effects of 1 months' additional progress in GPS. In light of the complexity of metacognition and considering that the ReflectED intervention is about an approach to learning and self-reflection, it is possible that it would be beneficial to consider longer term follow-up in similar evaluations in relation to pupil outcomes to ascertain if there are longer term effects.

We undertook an interaction test, which showed that there was a qualitative interaction between FSM status and treatment effect. Thus, the treatment effect was positive for the FSM children and negative for the non-FSM pupils (i.e. -1.85, 95% CI -4.52 to 0.82, p = 0.18). However, neither treatment effect was statistically significant.

Interpretation

The impact results of this RCT, evaluating the effectiveness of ReflectED, do not support the evidence in the Teaching and Learning Toolkit (EEF, 2021) on the effectiveness of metacognition and self-regulation strategies. Although the toolkit concludes the potential impact of these strategies is high, it does also point out that it may be difficult to realise that impact in practice. The evidence base for the toolkit conclusions was quite high—although a large number of studies were included in the synthesis, many of them were not independent evaluations. There is evidence in this evaluation from the IPE that could explain why any potential impact was not realised in this study in relation to variability in implementation and fidelity—this was discussed in full in the Compliance and Fidelity sections of the IPE.

This is a report of an effectiveness trial following the EEF-funded efficacy trial. The results from the current effectiveness trial partially support the results from the earlier cluster efficacy trial (randomisation by class) undertaken by the University of Manchester (Motteram *et al.,* 2016), in which a positive effect was found for mathematics but not for reading, although neither effect size was statistically significant. However, the confidence intervals around the efficacy trial effect sizes almost overlap with the effect sizes we found in our effectiveness trial. The efficacy trial obtained a high security rating of 4 padlocks and had a reasonable sample size of 70 classes with low attrition.

As was discussed in the introduction, metacognition is a 'fuzzy' concept (Wellman, 1985)—it is possible that the difficulty of realising the potential impact of metacognition and self-regulation strategies in practice could link to the challenges of defining a concept like metacognition. Metacognition has been described in different ways that are not always in agreement with each other (see Gascoine *et al.,* 2017 for further discussion of this). Kyriakides, *et al.* (2020) explained the importance of recognising the challenges of establishing what metacognitive practice is (in terms of pedagogy) and

how this often relies on self-reports from teachers, pupils, or both. Alongside this there is research (Branigan and Donaldson, 2020; Wall and Hall, 2016) emphasising the importance of the role of teachers in pupil metacognitive development. Not unlike the research literature that explores metacognition, data gathered in the IPE in relation to usual 'metacognition' practice at baseline is self-reported and it describes a variety of conceptualisations of metacognition among participating schools and individual teachers (control and intervention headteachers and classroom teachers). There are two key and interlinked issues to consider here when interpreting the results of this evaluation: first, the complexity of metacognition as a concept in theory and research (including the challenge of defining and describing it); and second, what this looks like in practice (not least how different schools and individual teachers describe and conceptualise it in practice). These two key issues are pertinent to interpreting the results of this evaluation and weighing up the extent to which the effectiveness of the ReflectED approach is influenced by the inherent complexity of metacognition itself.

There is a need to interrogate the links between how metacognition is defined, operationalised (pedagogically), and assessed in terms of considering the impact of a metacognition focused intervention on metacognitive skills (and how this is measured) and the impact on other cognitive, attainment focused outcomes as in this evaluation. Gascoine, *et al.,* (2017) drew on the work of Desoete (2008) in describing the importance of how metacognition is defined and 'tested' (or assessed) and what this means for what you get in terms of the outcomes of a study focused on metacognition. In this evaluation, using the JrMAI (Sperling, *et al.,* 2002; Sperling, *et al.,* 2012) there were no observed effects on metacognition. Given the complexity of measuring and assessing metacognition (e.g. Gascoine, *et al.,* 2017; Desoete, 2008) it is not unrealistic to reflect on the appropriateness of measuring metacognition in the context of this evaluation, and whether or not the gains described by the pupils and their teachers (in the qualitative data collection of the IPE) should be considered with more influence. Indeed, initial JrMAI research (Sperling, *et al.,* 2002, p. 53) noted that 'the use of achievement measures as indications of metacognitive knowledge or other self-regulatory constructs is unwarranted'. While Sperling, *et al.* (2002) are not saying that metacognition should not be measured, the complexity when combined with exploring metacognition in relation to academic achievement is clear. The inherent focus on academic achievement in this evaluation, and in the previously discussed wider research focused on metacognition means that it is of key importance to consider wider issues (including measuring) in the field of metacognition when interpreting the results of this evaluation.

Wang, *et al.* (2021) asserted the importance of understanding where the starting point is for pupil metacognitive skills and considering that perhaps they cannot 'reap the benefits of a growth mindset, without some basic levels of metacognitive skills' (p. 16). Links could be drawn here to this evaluation of ReflectED and the IPE evidence that suggest a whole range of previous teacher experience and teaching of metacognitive skills. Mindsets were a specific aspect of the ReflectED intervention that perhaps presented a barrier to impact (as per the impact evaluation) because there was variability in the baseline levels of metacognitive skills across the participating schools. Beyond this, growth and fixed mindsets (Dweck, 2016) are something that have been widely debated in the literature. In recent years, Dweck has cautioned over the risk of 'creat[ing] false growth-mindsets' (Dweck, 2015, p. 2) where fixed mindsets in relation to learning are 'banned', and emphasised the importance of not hiding gaps in achievement or learning but recognising them (via mindsets) to address them. The IPE presented evidence of teachers and pupils' views of the value of the ReflectED performance tag colours, which aligns with this approach (and caution from Dweck)—one where learning and gaps in it are visualised and teachers and pupils can work together to address these.

Whist it is clear in the literature that there is evidence to support the value of metacognition focused interventions for improving attainment in core subjects, including mathematics (Desoete and De Craene, 2019; Jones,*et al.,* 2020; Wang, *et al.,* 2021), it is also clear that this is complex and not without challenges. Metacognitive strategies or types of metacognitive instruction are most often not taught alone (the complexity and range of definitions interactions with other approaches would make this extremely difficult) and it is not uncomplicated to unpick interplay between these, nor to assess exactly what is having an impact on attainment where there is impact.

With similarity to ReflectED, the approach to metacognition explored by Wang, *et al.* (2021) included growth mindset. Unlike this study, the outcome focus was mathematics engagement rather than attainment; there is a link here to data gathered in the IPE where teachers described improvements specifically relating to mathematics and a perceived improvement in pupil's awareness of their learning process(es) in relation to mathematics in response to delivery of the ReflectED intervention and a greater level of resilience to learning. The teacher-described improvements were not reflected in the results of the impact evaluation; we have already considered the role that what we have described as the 'triple challenge' of this evaluation may have played. Although not focused on teacher perceived impacts, Jones, *et*

al. (2020) raised a question in their study about whether or not the metacognition focused aspect of their intervention ('MetaCogMed') (the other aspect being working memory training) was ineffective at promoting generalisation to subject focused tasks (e.g. mathematics, reading)—in order to explore this in relation to ReflectED further research would be required.

In summary, in interpreting the results of this evaluation, the multi-layered and interlinked challenges of evaluating an intervention such as ReflectED that we have considered in interpreting the results of this evaluation are:

1. ReflectED is a complex intervention with different components, that is delivered on a whole-school basis;

2. ReflectED is focused on a complex concept (metacognition) where there is long standing debate around defining, in both theory and practice; and

3. the challenges of evaluating the impact of ReflectED (a metacognition focused intervention) where the primary outcome measures are cognitive (i.e. outcomes in mathematics and reading) and not the non-cognitive (e.g. specific aspects of metacognitive skills that were developed in the ReflectED intervention).

Limitations and lessons learned

This evaluation had many strengths, including being designed as a large cluster RCT with in-depth IPE. Attrition at the school level was low, although the three schools, which withdrew before post-test, were all in the control group. Pupil-level attrition was also low at around 8% for each of the two primary outcomes. The IPE included multiple sources of data, which were integrated to address the research questions about implementation, fidelity, compliance, and so on.

There was evidence, in the IPE, of control school activity that included reflection on learning and metacognition focused activities—this was to be expected. As discussed in the background section there is extensive evidence (EEF, 2021) in support of the value of metacognition for academic progress—given this, it was to be expected that schools would be implementing a variety of metacognitive and self-regulation focused strategies. Therefore, it was an important part of the IPE to actively enquire as to what approaches to metacognition looked like across all of the participating schools. The impact of ReflectED, when considering the secondary JrMAI outcome for KS2 (no impact) also needs to be considered alongside the challenges and limitations of using self-reports to measure or assess metacognition.

There are limitations in relation to the use of the JrMAI (Sperling, *et al.,* 2002; Sperling, *et al.,* 2012) as a self-reported measure of metacognition. The challenges of self-reports have been explored in detail in the literature, and in similarly to the findings of a systematic review of ways of measuring or assessing metacognition in school-aged children (Gascoine, *et al.,* 2017), Wang, *et al.* (2021) note the value of considering behavioural and/or observational assessments of metacognition. Behavioural or observational assessments of metacognition in the future could be a useful way of unpicking the interactions between different aspects of the ReflectED intervention and exploring, which aspects specifically are having an impact on pupil metacognitive skills (related to the point above about exploring interactions between growth mindset and metacognitive skills):

> *Understanding the constraints and thresholds of how much or the specific type of metacognitive skill that is enough to boost growth mindset (and vice versa) and the parameters of these factors will further advance learning theories and inform practice in education. (Wang, et al., 2021, p. 17)*

Behavioural or observational assessments of metacognition in this study would have undoubtedly provided increased validity and reliability in relation to the assessment of metacognition, but on the other hand they would have increased the burden on participating schools, and it would have been a huge (and costly) undertaking to do this with an appropriate sample size in an evaluation of this size.

The limitations of the IPE focus on the challenge of synthesising and interpreting a large data set and explaining how the reality of collecting implementation and process data in school settings is complex. The number of school visits for the IPE (incorporating structured observations, teacher interviews, and pupil focus groups) is one example. Sometimes visits were cancelled at the last minute and other times we were unable to complete all aspects and therefore were not able to gather all the expected data (e.g. we might have been able to do a teacher interview and pupil focus group, but not a ReflectED lesson observation due to staff absence). In relation to research question 13, as explained in the IPE results, we were unable to collect a sample of reflections that had been completed and that would have facilitated analysis using the Moseley, *et al.* (2005) framework as planned. This deviation from the planned analysis of evidence

of metacognition in pupil reflections, completed as part of ReflectED, means that we cannot fully answer this research question with a high degree of validity because a different source of data (primarily from the self-reported post-intervention survey) provided the information for this research question rather than the evaluators being able to utilise the source of data as originally intended and independently assess a sample of pupil reflections for evidence of metacognition.

Future research and publications

In terms of future research, the evaluation has provided support for a model of training for ReflectED that is cascaded after selected practitioners (headteacher and one nominated lead practitioner in this evaluation) attend a launch training day in which they are introduced to the ReflectED intervention, have opportunities to see it in action, and an opportunity to be introduced to some of the key theories in relation to metacognition and self-regulation that underpin it. Although the impact evaluation did not show statistically significant effects on the primary outcome, IPE data around compliance and acceptance of a cascade training model like this could be useful for other educational evaluations.

The main challenge faced by schools, in terms of compliance, related to the completion of pupil reflections—further research is needed here in terms of the content and focus of reflections, how reflection is facilitated, and what frequency could support the future development of the ReflectED intervention and be a useful avenue through which to explore the impact of ReflectED on particular aspects of metacognition and change in metacognition over time. This would be justified given that previous research has shown positive links between metacognition and attainment *(*Akyol, *et al.,* 2010; Desoete and De Craene, 2019; Diginath and Büttner, 2008; Jones, *et al.,* 2020; Kuyper, *et al.,* 2000; Prins, *et al.,* 2006; Wang, *et al.,* 2021) in core curriculum subjects including mathematics and English.

Following the publication of this report we intend to publish a paper focused on the evaluation as a whole—exploring the multi-layered challenge of evaluating a complex, whole-school intervention, when it is focused on a complex concept like metacognition. We also intend to publish a paper that focuses on the qualitative analysis of survey data that asked the question: ‘What does metacognition look like?’

References

Akyol, G. Sungur, S. and Tekkaya, C. (2010). 'The Contribution of Cognitive and Metacognitive Strategy Use to Students' Science Achievement'. Educational Research and Evaluation, 16(1), 1–21.

Anders, J. Brown, C. Ehren, M. Greany, T. and Nelson, R. (2017). '*Evaluation of Complex Whole-School Interventions: Methodological and Practical Considerations'.* London, UK: Education Endowment Foundation (EEF). Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/eef-evaluation-reports-and-research-papers/methodological-research-and-innovations/evaluation-of-complex-whole-school-interventions (accessed 01 August 2022).

Bartsch, K. Horvath, K. and Estes, D. (2003). 'Young Children's Talk About Learning Events'. Cognitive Development, 18(2), 177–193.

Branigan, H. E. and Donaldson, D. I. (2020). 'Teachers Matter for Metacognition: Facilitating Metacognition in the Primary School Through Teacher-Pupil Interactions'. Thinking Skills and Creativity, 38, 100718.

Braun, V. and Clarke, V. (2021). 'One Size Fits All? What Counts as Quality Practice in (Reflexive) Thematic Analysis?' Qualitative Research in Psychology, 18(3), 328–352.

Data Protection Act (DPA). (2018). *Data Protection [online].* Available at: **https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted** [Accessed 30 August 2020].

Desautel, D. (2009). 'Becoming a Thinking Thinker: Metacognition, Self-Reflection, and Classroom Practice'. Teachers College Record, 111(8), 1997–2020.

Desoete, A. (2008). 'Multi-Method Assessment of Metacognitive Skills in Elementary School Children: How You Test is What You Get'. Metacognition and Learning, 3(3), 189–206.

Desoete, A. (2009). 'Metacognitive Prediction and Evaluation Skills and Mathematical Learning in Third-Grade Students'. Educational Research and Evaluation, 15(5), 435–446.

Desoete, A. and De Craene, B. (2019). 'Metacognition and Mathematics Education: An Overview'. ZDM – Mathematics Education, 51(4), 565–575.

Diginath, C. and Büttner, G. (2008). 'Components of Fostering Self-Regulated Learning Among Students. A Meta-Analysis on Intervention Studies at Primary and Secondary School Level'. Metacognition and Learning, 3(3), 231–264.

Dunn, G., Maracy, M. and Tomenson, B. (2005) 'Estimating Treatment Effects from Randomized Clinical Trials with Noncompliance and Loss to Follow-Up: The Role of Instrumental Variable Methods', Statistical Methods in Medical Research, 14 (4), pp. 369–395.

Dweck, C. (2015). 'Carol Dweck Revisits the Growth Mindset'. Education Week, 35(5), 20–24.

Dweck, C. (2016). 'What Having a "Growth Mindset" Actually Means'. Harvard Business Review*,* 13, 213–226.

EEF. (2015). *Intra-cluster correlation coefficients*. London: Education Endowment Foundation (EEF). Available at: **https://educationendowmentfoundation.org.uk/public/files/Evaluation/Writing_a_Protocol_or_SAP/ICC_2015.pdf** [Accessed 12 December 2022].

EEF. (2016). *Cost Evaluation Guidance*. London: Education Endowment Foundation (EEF). (No longer available online, as replaced with version of the guidance which is referenced and linked below.)

EEF. (2019). *Cost evaluation guidance for EEF evaluations [online].* Eduction Endowment Foundation (EEF). Available at: **https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/Cost_Evaluation_Guidance_2019.12.11.pdf** [Accessed 1 July 2022].

EEF. (2021). '*Teaching and Learning Toolkit'.* London, UK: Education Endowment Foundation (EEF). Available at: https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit (accessed 30 January 2022).

Fairhurst, C. and Roche, J. (2019). '*Statistical Analysis Plan: ReflectED*'. London, UK: Education Endowment Foundation (EEF). Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/reflected?utm_source=site&utm_medium=search&utm_campaign=site_search&search_term=ReflectEd (accessed 30 August 2020).

Flavell, J. H. 1976. 'Metacognitive Aspects of Problem Solving'. *In* L. B. Resnick ed. *The Nature of Intelligence*. Hillsdale, New Jersey, Lawrence Erlbaum Associates, Inc. pp. 231–235.

Gascoine, L. 2016. *'Investigating Developmental Trends in Metacognitive Knowledge with School-Aged Children Using Pupil Views Templates'.* Doctoral Thesis. Durham, UK: School of Education, Durham University.

Gascoine, L. Higgins, S. and Wall, K. (2017). 'The Assessment of Metacognition in Children Aged 4–16 Years: A Systematic Review'. Review of Education, 5(1), 3–57.

Gascoine, L. Torgerson, D. J. Tracey, L. C. Bell, K. Fairhurst, C. Elliott, L. et al. (2018a). '*Evaluation of ReflectED (Trial Protocol)'*. Available at: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/ReflectED_protocol_18.09.06__updated.pdf (accessed 30 January 2020).

Gascoine, L. Torgerson, D. J. Tracey, L. C. Bell, K. Fairhurst, C. Elliott, L. et al. (2018b). '*Evaluation of ReflectED (Trial Protocol, Revised)'*. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/projects/ReflectED_protocol_18.09.06__updated.pdf?v=1671010698 (accessed 30 January 2020).

General Data Protection Regulation. (2016). '*Art. 6 GDPR: Lawfulness of Processing*'. Available at: https://gdpr-info.eu/art-6-gdpr/ (accessed 30 November 2021).

Humphrey, N. Lendrum, A. Ashworth, E. Frearson, K. Buck, R. and Kerr, K. (2016). '*Implementation and Process Evaluation (IPE) for Interventions in Educational Settings: An Introductory Handbook'.* London, UK: Education Endowment Foundation (EEF). Available at: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Handbook.pdf (accessed 01 September 2017).

Jones, J. S. Milton, F. Mostazir, M. and Adlam, A. R. (2020). 'The Academic Outcomes of Working Memory and Metacognitive Strategy Training in Children: A Double-Blind Randomized Controlled Trial'. Developmental Science, 23(4), e12870.

Kuhn, D. 1999. 'Metacognitive Development'. *In* L. Balter and C. S. Tamis-LeMonda eds. *Child Psychology. A Handbook of Contemporary Issues*. Philadelphia: Psychology Press. pp. 259–286.

Kuyper, H. van der Werf, M. P. C. and Lubbers, M. J. (2000). 'Motivation, Meta-Cognition and Self-Regulation as Predictors of Long Term Educational Attainment'. Educational Research and Evaluation, 6(3), 181–205.

Kyriakides, L. Anthimou, M. and Panayiotou, A. (2020). 'Searching for the Impact of Teacher Behavior on Promoting Students' Cognitive and Metacognitive Skills'. Studies in Educational Evaluation, 64, 100810.

Leutwyler, B. (2009). 'Metacognitive Learning Strategies: Differential Development Patterns in High School'. Metacognition and Learning, 4(2), 111–123.

Limitation Act. (1980). *Limitation Act 1980: 1980 Chapter 58 [online]*. Available at: **https://www.legislation.gov.uk/ukpga/1980/58** [Accessed 30 August 2020].

Moore, G. F. Audrey, S. Barker, M. Bond, L. Bonell, C. Hardeman, W. et al. (2015). 'Process Evaluation of Complex Interventions: Medical Research Council Guidance'. BMJ (Clinical Research Ed.), 350, h1258.

Moseley, D. Baumfield, V. Elliott, J. Higgins, S. Miller, J. Newton, D. P. et al. (2005). '*Frameworks for Thinking: A Handbook for Teaching and Learning'*. Cambridge, UK: Cambridge University Press.

Motteram, G. Choudry, S. Kalambouka, A. Hutcheson, G. and Barton, A. (2016). '*ReflectED: Evaluation Report and Executive Summary'*. London, UK: Education Endowment Foundation (EEF). Available at: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/EEF_Project_Report_ReflectED.pdf. [Accessed 01 September 2017]

NVivo. Version (2018). *'NVivo Qualitative Data Analysis Software'*. Victoria, Melbourne: QSR International Pty Ltd.

Oakley, A. Strange, V. Bonell, C. Allen, E. Stephenson, J. and RIPPLE Study Team. (2006). 'Process Evaluation in Randomised Controlled Trials of Complex Interventions'. BMJ (Clinical Research Ed.), 332(7538), 413–416.

Prins, F. J. Veenman, M. V. J. and Elshout, J. (2006). 'The Impact of Intellectual Ability and Metacognition on Learning: New Support for the Threshold of Problematicity Theory'. Learning and Instruction, 16, 374–387.

Quigley, A. Muijs, D. and Stringer, E. (2018). '*Metacognition and Self-Regulated Learning: Guidance Report'.* London, UK: Education Endowment Foundation (EEF).

Records Management: code of practice for health and social care (2016). *Records managament code of practice [online]*. Available at: **https://www.gov.uk/government/publications/records-management-code-of-practice-for-health-and-social-care** [Accessed 30 August 2020].

ReflectED. (2021). '*About ReflectED'.* Available at: https://www.reflectedlearning.org.uk/about-reflected/ (accessed 01 July 2021).

Saghaei, M. and Saghaei, S. (2011). 'Implementation of an Open-Source Customizable Minimization Program for Allocation of Patients to Parallel Groups in Clinical Trials'. Journal of Biomedical Science and Engineering, 4, 734–739.

Sperling, R. A. Howard, B. C. Miller, L. A. and Murphy, C. (2002). 'Measures of Children's Knowledge and Regulation of Cognition'. Contemporary Educational Psychology, 27(1), 51–79.

Sperling, R. A. Richmond, A. S. Ramsay, C. M. and Klapp, M. (2012). 'The Measurement and Predictive Ability of Metacognition in Middle School Learners'. The Journal of Educational Research, 105(1), 1–7.

Thomas, D. R. (2006). 'A General Inductive Approach for Analyzing Qualitative Evaluation Data'. American Journal of Evaluation, 27(2), 237–246.

Veenman, M. V. J. Kok, R. and Blöte, A. W. (2005). 'The Relation Between Intellectual and Metacognitive Skills in Early Adolescence'. Instructional Science, 33(3), 193–211.

Veenman, M. V. J. Wilhelm, P. and Beishuizen, J. J. (2004). 'The Relation Between Intellectual and Metacognitive Skills From a Developmental Perspective'. Learning and Instruction, 14(1), 89–109.

Wall, K. (2008). 'Understanding Metacognition Through the Use of Pupil Views Templates: Pupil Views of Learning to Learn'. Thinking Skills and Creativity, 3(1), 23–33.

Wall, K. and Hall, E. (2016). 'Teachers as Metacognitive Role Models'. European Journal of Teacher Education, 39(4), 403–418.

Wall, K. Higgins, S. and Packard, E. (2007). '*Talking About Learning: Using Templates to Find Out Pupil's Views'*. Devon: Southgate Publishers.

Wall, K. Higgins, S. Remedios, R. Rafferty, V. and Tiplady, L. (2013). 'Comparing Analysis Frames for Visual Data Sets: Using Pupil Views Templates to Explore Perspectives of Learning'. Journal of Mixed Methods Research, 7(1), 22–42.

Wang, M. T. Zepeda, C. D. Qin, X. Del Toro, J. and Binning, K. R. (2021). 'More Than Growth Mindset: Individual and Interactive Links Among Socioeconomically Disadvantaged Adolescents' Ability Mindsets, Metacognitive Skills, and Math Engagement'. Child Development, 92(5), e957–e976.

Wellman, H. M. (1985). 'The Origins of Metacognition'. Metacognition, Cognition, and Human Performance, 1, 1–31.

Whitebread, D. Coltman, P. Pasternak, D. P. Sangster, C. Grau, V. Bingham, S. et al. (2009). 'The Development of Two Observational Tools for Assessing Metacognition and Self-Regulated Learning in Young Children'. Metacognition and Learning, 4, 63–85.

Yeager, D. S. Hanselman, P. Walton, G. M. Murray, J. S. Crosnoe, R. Muller, C. et al. (2019). 'A National Experiment Reveals Where a Growth Mindset Improves Achievement'. Nature, 573(7774), 364–369.

Appendix A: EEF cost rating

Appendix Figure 1: 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: Key Stage 2 statutory assessment in maths

Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	MDES	Attrition					
5	Randomised design	<= 0.2	0-10%	5				5
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%					
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 [0]		
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	Threat to internal validity?	Comments
Threat 1: Confounding	MODERATE	Well-designed RCT design with randomisation conducted by evaluation team. There are some imbalances between arms in the pre-test measure (.05-.1 SD), controlled for in analysis.
Threat 2: Concurrent Interventions	LOW	Concurrent interventions are explored via the IPE and similar levels of practice related to metacognition across arms. There were similar levels of additional interventions aimed at improving maths/literacy, although control schools did report slightly higher incidence of whole school maths interventions (45%) compared to intervention schools (31%) at the end of the evaluation.
Threat 3: Experimental effects	LOW	Conditions were monitored through usual practice. There is report of one control school implementing ReflectED but no other evidence that control schools increased their metacognitive practices (beyond usual practice) during the intervention period.
Threat 4: Implementation fidelity	MODERATE	Fidelity and compliance were well-defined and aligned with the logic model. Lesson plans seem to have been implemented by teachers, but the number of reflections completed/recorded was below 2-week threshold, and compliance data were not well recorded by schools.
Threat 5: Missing Data	Low	Missing data less than 10% and analysis based on multiple imputation was similar to complete case analysis
Threat 6: Measurement of Outcomes	LOW	Primary outcome measurement was appropriate. Reliable valid test data was used from NPD.
Threat 7: Selective reporting	LOW	Analyses were pre-specified within the published protocol and SAP. The trial was registered on ISRCTN. The evaluation followed the protocol closely.

- **Initial padlock score:** 5 padlocks; well-designed two-arm cluster randomised trial; MDES of .18 at randomisation and less than 9% attrition experienced.
- **Reason for adjustment for threats to validity**: 0 padlocks lost. Two moderate threats to validity have been identified for the trial. First, there were a few marginal imbalances greater than .05 SD in pupil prior attainment between arms that pose a moderate confounding risk. Second, whilst there is indication from the process evaluation that teachers implemented ReflectED lessons plans, this is not systematically captured given the high level of missing compliance data. Overall as direction of bias is uncertain, no further deduction in padlocks.
- **Final padlock score:** 5 padlocks

OUTCOME: Key Stage 2 statutory assessment in reading

Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	MDES	Attrition					
5	Randomised design	<= 0.2	0-10%	5				5
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%					
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 [0]		
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	Threat to internal validity?	Comments
Threat 1: Confounding	MODERATE	Well-designed RCT design with randomisation conducted by evaluation team. There are some imbalances between arms in the pre-test measure (.05-.1 SD), controlled for in analysis.
Threat 2: Concurrent Interventions	LOW	Concurrent interventions are explored via the IPE and similar levels of practice related to metacognition across arms. There were similar levels of additional interventions aimed at improving maths/literacy, although control schools did report slightly higher incidence of whole school maths interventions (45%) compared to intervention schools (31%) at the end of the evaluation.
Threat 3: Experimental effects	LOW	Conditions were monitored through usual practice. There is report of one control school implementing ReflectED but no other evidence that control schools increased their metacognitive practices (beyond usual practice) during the intervention period.
Threat 4: Implementation fidelity	MODERATE	Fidelity and compliance were well-defined and aligned with the logic model. Lesson plans seem to have been implemented by teachers, but the number of reflections completed/recorded was below 2-week threshold, and compliance data were not well recorded by schools.

Threat 5: Missing Data	Low	Missing data less than 10% and analysis based on multiple imputation was similar to complete case analysis.
Threat 6: Measurement of Outcomes	LOW	Primary outcome measurement was appropriate. Reliable valid test data was used from NPD.
Threat 7: Selective reporting	LOW	Analyses were pre-specified within the published protocol and SAP. The trial was registered on ISRCTN. The evaluation followed the protocol closely.

- **Initial padlock score:** 5 padlocks; well-designed two-arm cluster randomised trial; MDES of .18 at randomisation and less than 9% attrition experienced.
- **Reason for adjustment for threats to validity**: 0 padlocks lost. Two moderate threats to validity have been identified for the trial. First, there were a few marginal imbalances greater than .05 SD in pupil prior attainment between arms that pose a moderate confounding risk. Second, whilst there is indication from the process evaluation that teachers implemented ReflectED lessons plans, this is not systematically captured given the high level of missing compliance data. Overall as direction of bias is uncertain, no further deduction in padlocks.
- **Final padlock score:** 5 padlocks

Appendix C: Changes since the previous evaluation

Appendix Table 1: Changes since the previous evaluation

	Feature	Efficacy to effectiveness stage
Intervention	Intervention content	In the efficacy trial participating pupils were expected to complete a minimum of one reflection per week, in the effectiveness trial the minimum requirement was two reflections per week. In the effectiveness trial ReflectED was delivered as whole school intervention, lesson plans (one per week) were provided for each year group in the participating schools from EYFS up to Year 6.
	Delivery model	The training approach differed in that in the effectiveness trial a cascaded model of training was used where the head teacher and nominated lead practitioner from each school attended launch training held at Rosendale School and then were tasked with cascading this training to the rest of the teaching staff in their school. In the efficacy trial the focus was on the delivery of ReflectED to year 5 pupils, in the effectiveness trial ReflectED was delivered as a whole school intervention. The focus on Year 1 (KS1 cohort) and Year 5 (KS2 cohort) at the recruitment stage and then the same pupils when they were in Year 2 and Year 6 respectively was driven by using NPD data for outcomes to minimise burden on participating schools.
	Intervention duration	ReflectED was delivered over one academic year in the efficacy trial. In the effectiveness trial the intervention was delivered over 5 school terms, spanning two academic years.
Evaluation	Eligibility criteria	In the efficacy trial only schools with two form entry and above were eligible to participate, in the effectiveness trial one form entry was the minimum threshold for eligibility in terms of school size.
	Level of randomisation	Efficacy trial – class level randomisation; Effectiveness trial – school level randomisation
	Outcomes and baseline	**Outcomes** Efficacy trial: Primary outcome = age standardised mathematics score (InCAS, CEM) Secondary outcomes = standardised reading scores, standardised attitude to mathematics scores, standardised attitude to reading scores (InCAS, CEM) Effectiveness trial (current): Primary outcomes (KS2 cohort) = (a) KS1 Mathematics score (NPD), (b) KS2 Reading score (NPD) Secondary outcomes (KS2 cohort) = c) Grammar, punctuation, and spelling (GPS) (NPD), d) Junior Metacognitive Awareness Inventory (JrMAI) (collected by YTU) Secondary outcomes (KS1 cohort) = e) KS1 mathematics score (raw data obtained directly from schools), f) KS1 reading scores (raw data obtained directly from schools) **Baseline:** Efficacy trial: KS1 Mathematics score (NPD) Effectiveness trial (current): Primary outcomes (KS2 cohort) = a) KS1 mathematics score (NPD), b) KS1 reading score (NPD) Secondary outcomes (KS2 cohort) = c) KS1 writing attainment point score (NPD), d) pre-test JrMAI collected prior to randomisation ○ Secondary outcomes (KS1 cohort) = e) and f) Early Years Foundation Stage Profile (EYFSP)
	Control condition	In this evaluation, control groups were provided with an opportunity to pick from a selection of pupil premium focused CPD.

Appendix D: Effect size estimation

Appendix Table 2: Effect size estimation

			Intervention group		Control group		
Outcome	Unadjusted differences in means (95% CI)	Adjusted differences in means (97.5%/95% CI)	n (missing)	Variance of outcome	n (missing)	Variance of outcome	Pooled variance
KS2 Mathematics	-1.62 (-3.07, -0.17)	-1.14 (97.5% CI -4.36, 2.08)	2217 (64)	584.85	2081 (164)	586.71	588.69
KS2 Reading	-0.12 (-0.73, 0.49)	0.06 (97.5% CI -1.06, 1.17)	2209 (72)	104.97	2088 (157)	103.39	97.85
KS2 GPS	-1.20 (-2.06, -0.34)	-1.00 (95% CI -2.56, 0.56)	2210 (71)	216.84	2087 (158)	198.28	209.08
JrMAI	-0.11 (-0.31, 0.09)	-0.13 (95% CI -0.49, 0.22)	1942 (339)	8.86	1758 (487)	9.50	9.15
KS1 Mathematics	-0.47 (-1.44, 0.50)	-0.16 (95% CI -1.98, 1.66)	1909 (412)	255.75	1841 (428)	198.42	225.44
KS1 English	-0.15 (-0.87, 0.57)	0.35 (95% CI -0.87, 1.57)	1896 (425)	129.80	1751 (518)	198.42	111.92
KS2 Mathematics FSM subgroup	1.64 (-1.12, 4.40)	1.30 (95% CI 2.57, 5.16)	717 (35)	664.82	642 (69)	674.84	669.33
KS2 Reading FSM subgroup	0.59 (-0.57, 1.75)	0.47 (95% CI -0.96, 1.90)	714 (38)	120.50	646 (65)	115.44	118.21

Further appendices

Please see accompanying document ‘Technical notes’.

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

Education
Endowment
Foundation

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

@EducEndowFoundn

Facebook.com/EducEndowFoundn