Education Endowment Foundation

Dialogic Teaching

Evaluation report and executive summary

July 2017

Independent evaluators:

Professor Tim Jay, Ben Willis, Dr Peter Thomas, Dr Roberta Taylor, Dr Nick Moore, Professor Cathy Burnett, Professor Guy Merchant, Anna Stevens

Sheffield Hallam University

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

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

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

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

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

Danielle Mason
Head of Research and Publications
Education Endowment Foundation
9th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

p: 020 7802 1679
e: danielle.mason@eefoundation.org.uk
w: www.educationendowmentfoundation.org.uk

About the evaluator

The project was independently evaluated by a team from Sheffield Hallam University: Professor Tim Jay, Ben Willis, Dr Peter Thomas, Dr Roberta Taylor, Dr Nick Moore, Professor Cathy Burnett, Professor Guy Merchant, Anna Stevens

The lead evaluator was Professor Tim Jay.

Contact details:

Professor Tim Jay
Centre for Education and Inclusion Research, Sheffield Hallam University
City Campus, Howard Street, Sheffield S1 1WB, UK
Email: T.Jay@shu.ac.uk

Tel: 0114 225 4850

Contents

Executive summary

The project

The aim of the intervention was to raise levels of engagement and attainment across English, maths, and science in primary schools by improving the quality of teacher and pupil talk in the classroom. The approach, termed 'dialogic teaching', emphasises dialogue through which pupils learn to reason, discuss, argue, and explain in order to develop their higher order thinking as well as their articulacy. The intervention was developed and delivered by a team from the Cambridge Primary Review Trust (CPRT) and the University of York. Year 5 teachers in 38 schools, and a teacher mentor from each school, received resources and training from the delivery team, and then implemented the intervention over the course of the autumn and spring terms in the 2015/2016 school year. Following the intervention, pupils were tested in English, mathematics, and science. This efficacy trial compared the 38 schools (2,492 pupils) in which the intervention took place with 38 control schools (2,466 pupils). During the intervention, the evaluation team also carried out a survey and interviews with a sample of teachers, mentors, and heads, plus case-study visits to three intervention schools.

Key conclusions	
1.	Children in Dialogic Teaching schools made two additional months' progress in English and science, and one additional month's progress in maths, compared to children in control schools, on average. The three padlock security rating means we are moderately confident that this difference was due to the intervention and not to other factors.
2.	Children eligible for free school meals (FSM) made two additional months' progress in English, science, and maths compared to FSM children in control schools. The smaller number of FSM pupils in the trial limits the security of this result.
3.	The intervention was highly regarded by headteachers, mentors, and teachers who thought that the Dialogic Teaching approach had positive effects on pupil confidence and engagement.
4.	The majority of participating teachers felt that it would take longer than two terms to fully embed a Dialogic Teaching approach in their classrooms. It could therefore be valuable to test the impact of the intervention over a longer period.
5.	This intervention requires teachers to change classroom talk across the curriculum, supported by training, handbooks, video, and regular review meetings with mentors. Future research could aim to differentiate the effects of these different elements.

Security rating: 

EEF security rating

These findings have moderate security. The security rating of the trial indicates how confident we can be that any additional progress experienced by the children receiving the intervention was due to the intervention and not any other factors. This was an efficacy trial which tested whether the intervention can work under developer-led conditions.

This was a well-designed randomised controlled trial. Twenty-one percent of pupils were not included in the final analysis, primarily because seven out of thirty-eight schools in the intervention group failed to provide post-test data. However, the pupils who received the intervention were similar to the pupils in the comparison group on the characteristics that were measured. There is some evidence that schools implemented the intervention to varying extents, but not such that there was a threat to validity.

Additional findings

The process evaluation showed that the Dialogic Teaching approach was highly valued by participating schools. Teachers reported positive effects on pupil engagement and confidence. However, some

schools also found the approach very challenging to implement within the two terms that this project lasted (autumn and spring terms 2015/2016). In fact, teachers felt that the impact evaluation was unlikely to show a positive effect on attainment for this reason. This means it is possible that the effect sizes observed in this evaluation are underestimates of the potential impact.

A follow-up to the impact evaluation will be carried out and reported in an addendum to this report, due to be published in 2018. The follow-up will test for an effect on pupil scores in Key Stage 2 national tests in English and mathematics.

Separate from this evaluation, the project team undertook analysis of video data from treatment and control schools in order to compare changes in teacher and pupil talk over time, supplemented by interviews with teachers, mentors, and heads. Interim findings are included in the process evaluation section, and complete findings will be published separately following this report.

Cost

The cost has been estimated as £52 per pupil per year. In addition to this financial cost, up to 17 days per year of teaching cover could be required per school. This includes 11 days for the training (3 days per teacher, 4 days for the mentor, and 1 day for the headteacher), and 6 days for the planning and review meetings within school. In this project, this time was mostly met from schools' existing cover and planning arrangements.

Table 1: Summary of impact on primary outcomes

Subject area	Group	Effect size (95% confidence interval)	Estimated months' progress	EEF security rating	P value	EEF cost rating
Maths	Treatment vs. control	0.09 (-0.04, 0.20)	1	🔒🔒🔒🔓🔓	0.19	**£**££££
	Treatment vs. control (FSM only)	0.16 (0.01, 0.30)	2	n/a	0.03	**£**££££
Science	Treatment vs. control	0.12 (0.01, 0.23)	2	🔒🔒🔒🔓🔓	0.04	**£**££££
	Treatment vs. control (FSM only)	0.11 (-0.04, 0.26)	2	n/a	0.14	**£**££££
English	Treatment vs. control	0.15 (0.00, 0.30)	2	🔒🔒🔒🔓🔓	0.05	**£**££££
	Treatment vs. control (FSM only)	0.12 (-0.07, 0.30)	2	n/a	0.21	**£**££££

Introduction

Intervention

The Dialogic Teaching intervention is designed to improve the quality of classroom talk as a means of increasing pupils' engagement, learning, and attainment, particularly those from disadvantaged backgrounds. The programme builds on the dialogic teaching approach developed by Alexander (2015) and successfully implemented in several local authorities (Alexander 2003, 2005a, 2005b). The programme evaluated here was piloted in Barking and Dagenham in 2014/2015. It emphasises dialogue through which pupils learn to reason, discuss, argue, and explain, in order to develop their higher order thinking and articulacy.

The programme as evaluated here included training for teachers, ongoing in-school monitoring and support, a pack containing study and reference materials, and a development and mentoring manual. The intervention entailed a structured programme of 11 cycles arranged in two school terms, preceded by induction and training days. The core strategies were:

- mentoring;
- video and audio recording for self-evaluation and development; and
- an iterative process of target-setting, implementation, recording, and review.

Schools were provided with all necessary equipment for video and audio recording—an essential parts of the process—and two days of cover per teacher was paid for time when participating teachers were away from their classrooms. The intervention was delivered by Cambridge Primary Review Trust and the University of York, with training led by Robin Alexander.

Although there is strong evidence that the dialogic teaching approach can improve motivation, engagement, participation, and reasoning, there has not yet been a U.K. randomised controlled trial to assess its effectiveness in raising attainment. The following, more detailed, description of the intervention follows the 'Template for Intervention Description and Replication' (Hoffman *et al*., 2014).

1. Brief name

Dialogic Teaching.

2. Why (rationale/theory)

The intervention aimed:

- to maximise the quality and educational impact of classroom talk, building on prior work on dialogic teaching and on international evidence;
- to encourage a classroom culture that engages pupils in the task in hand and retains their attention and interest;
- to meet, but also go beyond, the requirements for spoken language in the national curriculum, giving particular attention to those kinds of talk through which pupils learn to reason, explain, justify, argue, speculate, evaluate, and in other ways think for themselves;
- to advance this higher-order talk across the curriculum, but devote particular attention to it in the teaching of English, mathematics, and science; and
- to raise pupils' standards of attainment in literacy, numeracy, and science above the levels that teaching without such an intervention is likely to achieve.

3. Who (recipients)

All pupils in Year 5 classes from eligible schools (> 20% eligibility for FSM). The Dialogic Teaching programme comprises a combination of direct induction, training, and plenary sessions led by the

delivery team with Year 5 teachers, a nominated school mentor, and the headteacher, alongside materials and resources and ongoing in-school monitoring and support.

4. What (materials)

Participating schools receive a range of resources from the delivery team to foreground the principles of dialogic teaching and to support the professional development of participants and their colleagues. In addition, schools receive audio-visual recording equipment to facilitate the recording of classroom talk—a vital element of the project's strategy.

Every participating teacher, mentor, and school head was provided by the project team with the following resources:

- The CPRT/University of York Dialogic Teaching Project, trial stage 2015/2016, 'Handbook for schools'. The handbook sets out the programme's aims, rationale, and strategies, specifies in detail the programme to be followed over the two terms, cycle by cycle, and provides extensive lesson transcript material to exemplify the various repertoires of teacher and pupil talk to which the intervention is directed.

- Separate booklet containing monitoring forms for each planning/review cycle, for completion during mentor meetings.

- Alexander, R. J. (2015) *Towards Dialogic Teaching: rethinking classroom talk*, (4th edition, 2015 reprint), York: Dialogos. This presents the approach and evidence on which the project is chiefly based, and lists additional professional sources and resources.

- Michaels, S. and O'Connor, C. (2012) *Talk Science Primer*, Boston MA: TERC. This book details teacher talk moves through which pupil talk can be extended and built upon in one subject, science.

- Alexander, R. J. (2015) 'Dialogic Teaching Repertoires': a laminated card summarising, from the two publications above, the talk repertoires which the project aims to explore, foster, and improve.

All participants had access to materials on the project website, including most of the publications above and two specially-prepared video presentations shown at the induction sessions:

- Dialogic Teaching. This contains a sequence of clips of teachers and pupils talking, with commentary. It is structured into sections dealing with dialogic teaching repertoires and moves.

- Video Recording in Classrooms. This provides basic advice on how to make good quality video and audio recordings for use during the project.

In addition to the above, each mentor received a copy of:

- Lefstein, A. and Snell, J. (2015) *Better than Best Practice: developing teaching and learning through dialogue*, Routledge. As well as its combination of commentary and transcribed lesson extracts, this gives access to a linked website containing video recordings of the lessons in question.

Finally, for the recording of classroom talk every school was provided with the following project audio-video equipment:

- Panasonic HC-W570EB-K Full HD Camcorder with Twin Camera
- Olympus VN-732PC 4Gb Digital Voice Recorder
- Hama Start 61 Tripod
- Transcend 64 GB Premium SDXC Class 10 Memory Card
- Batteries
- Camera case

5. What (procedures)

Year 5 teachers were asked to run 11 planning/review/refocusing cycles with their classes, six in phase 1 ('Expanding repertoires') and five in phase 2 ('Advancing dialogue')—meaning the entire intervention programme lasts for 20 weeks across two terms. During phase 1 ('Expanding repertoires'), the focus was on increasing and enhancing the different kinds of talk used by teachers and pupils. During phase 2 ('Advancing dialogue') the focus was on applying the repertoires within a six-week programme in English, mathematics, science, and one non-core subject. Each cycle had two foci—a 'directed' and 'responsive' focus. The directed focus was one that all teachers were expected to engage with during a given cycle, while the responsive focus represented an opportunity for teachers to adapt and develop the approach to suit the context of their own classroom. The specific requirements for each cycle are clearly set out in the project handbook but broadly speaking the project is incremental and progressive in that each cycle was designed to build on the cycles that went before it. Initially, ground rules for talk are established (for example, listen carefully, respect others' ideas, don't interrupt), then different teacher and pupil talk repertoires (for example, questioning, exposition, feedback, probing and building on pupil contributions, expanding pupil learning talk) are introduced and refined with a view to applying them in varied contexts (such as whole-class teaching, teacher-led small-group discussion, pupil-led small-group discussion, teacher-pupil one-to-one, and paired pupil to pupil). The intention of the project is that these repertoires should be applied across the curriculum, though within this project there was a particular focus on English, maths, and science.

6. Who (implementers)

The Dialogic Teaching approach is designed to be delivered by Year 5 class teachers within their regular lessons.

7. How (mode of delivery)

Year 5 teachers employ Dialogic Teaching as part of the normal Year 5 class timetable.

8. Where (location of the intervention)

The intervention took place in pupils' usual classrooms.

9. When and how much (dosage)

The principles of Dialogic Teaching are intended to inform lesson delivery across the curriculum, with a particular focus on English, maths, and science. The programme runs for 20 weeks across two terms.

10. Tailoring

Delivery of the Dialogic Teaching project is based on eleven plan/review/refocus cycles. Each cycle has both 'directed' and 'responsive' foci as detailed in the Handbook for Schools (see point 5 above). The directed focus represents the compulsory core intended to ensure consistency and progression whereas the responsive focus allows additional opportunity for class teachers or mentors to attend to other aspects of talk as an extension or addition to the directed focus.

11. How well (planned)

In addition to the eleven plan/review/refocus cycles, the Dialogic Teaching project combines various additional strategies to help support and maximise implementation effectiveness. These include:

Training, induction and plenaries: a series of sessions led by the delivery team to help ensure understanding of the project aims and expectations across all participants.

- July 2015—induction day for teachers, mentors and heads;
- September 2015—mentor training day;
- December 2015—plenary day for mentors; and
- May 2016—plenary day for teachers, mentors, and heads.

Video and audio recording: schools are provided with all necessary equipment for video and audio recording lessons. Recordings are used as baselines for future development and later comparison forming a key component of mentoring sessions.

Mentoring: each school appointed a mentor to support Year 5 teachers in planning and reviewing their project related activities. The mentoring relationship with teachers was intended to be one of peers, where professional learning is mutual and encourages open and non-judgemental discussion. Mentors were required to organise, for each of the eleven cycles, sessions for planning and review, working with their mentees either singly or jointly. Most opted to combine end-of-cycle reviews with planning of the next cycle, and for these purposes they usually brought the Year 5 teachers together so that they could share experiences and learn from each other. Playing and discussing video and audio footage from the cycle under review was an essential part of the process.

Professional study and discussion: participants are encouraged to read and discuss the ideas and theory on which the project is based. A comprehensive range of resources and materials are provided to participants as outlined above (section 4).

12. How well (actual): evidence of implementation variability

Aside from the early withdrawal of six schools shortly after randomisation, the process evaluation found there to be a high level of fidelity in relation to attendance of events led by the delivery team, broad adherence to the eleven planning/review/refocussing cycle format, and undertaking of mentor meetings. However, there was greater variability in terms of which resources and materials were engaged with, the duration of mentor meetings, and the extent to which Dialogic Teaching approaches were routinely applied within lessons across the full curriculum.

Background

The dialogic teaching approach has been strongly influenced by the work of Vygotsky and Bruner, two researchers with a focus on social and cultural aspects of learning and education. There are also strong links with Bakhtin, who may have first coined the term 'dialogism'.

Dialogic teaching can be placed in opposition to 'monologic' talk, which often characterises classroom interaction (Lyle, 2008). Monologic talk is that which is dominated by the teacher, and which generally

only includes children as part of an 'Initiation-Response-Feedback' (IRF) pattern. The IRF pattern can be seen in the classroom when a teacher asks a closed question (initiation) of a particular child, then the child provides an answer (response), then finally the teacher gives feedback for that answer (feedback). The IRF pattern may be repeated several times during a period of whole-class teaching. The IRF pattern, and monologic teaching in general, has been criticised for limiting the amount of talk with which pupils can meaningfully engage (Mercer, 1995). Dialogic teaching, on the other hand, involves talk that is cumulative, supportive, reciprocal, collective, and purposeful (Alexander, 2008). So, questions are structured to provoke thoughtful answers; answers prompt further questions and are seen as the building blocks of dialogue rather than its terminal point; and individual teacher-pupil and pupil-pupil exchanges are chained into coherent lines of enquiry rather than left disconnected (Alexander, 2008).

Alexander (2017) states that: "There are a number of related approaches that involve a focus on the development of classroom talk to promote learning. Though they share a commitment to raising the profile and power of classroom talk, and are often grouped under the umbrella terms 'dialogue' and 'dialogic', there are some important differences among them. Some of them focus more on the teacher's talk (for example Wragg and Brown, 1993, 2001) and some on the pupils' talk (such as Dawes, Mercer and Wegerif, 2004). Others, including the approach evaluated here, attend equally to both teacher and pupil talk, and to the relationship between them. In differentiating these various pedagogical approaches Lefstein and Snell (2014) show how they vary, not just in respect of strategy, but also in the way they reflect contrasting notions of dialogue's nature and purposes. In parallel, Alexander (2001, 2008) draws on his transnational and cross-cultural research to show how approaches to classroom talk are shaped by culturally-embedded stances on teaching more broadly conceived, which he differentiates as 'transmission', 'initiation', 'negotiation', 'facilitation', and 'acceleration'. The framework of Lefstein and Snell (2014) is helpful in distinguishing between different approaches. They identify four: dialogically organised instruction (Nystrand, 1997, 2006), exploratory talk (Mercer 2000, Mercer and Littleton, 2007), accountable talk (Resnick, Michaels and O'Connor, 2010), and dialogic teaching (Alexander, 2001, 2008)."1

Evidence for the effectiveness of dialogic teaching as a pedagogical approach can be found in research studies that have either made use of observation methods, or have involved small-scale intervention. Resnick, Asterhan and Clarke (2015) brings together a set of studies that provide a broad base of evidence for the effectiveness of structured dialogic teaching approaches in raising pupil attainment. Observational and quasi-experimental studies have been carried out that have explored the effects of implementing dialogic teaching in mathematics (Mercer and Sams, 2006), science (Mercer, Dawes, Wegerif and Sams, 2004; Mercer, Dawes and Staarman, 2009), and literacy (Reznitskaya, 2012). Such studies have shown interventions based on a dialogic teaching approach to be effective in increasing the quantity and quality of classroom talk, and in raising attainment. For example, Mercer and Sams (2006) report an evaluation of the Thinking Together intervention for mathematics learning in Year 5. This intervention led to substantial changes in classroom practice (seven teachers undertook the intervention, with 196 pupils), and significant gains in mathematics scores, with an effect size of +0.59. The Thinking Together intervention consisted of twelve lessons focusing on data handling, properties of numbers, and number sequences. Mercer and Sams (2006) is a good example of the evaluation that has been carried out on this approach to date: although positive effects were observed, the sample size was small, and the intervention focused on a particular section of one curriculum subject.

The Dialogic Teaching approach has influenced the development of the Philosophy for Children programme, which has been evaluated using a randomised controlled trial in a previous EEF project (Gorard, Siddiqui and Huat See, 2016). Children received, on average, one lesson per week of the Philosophy for Children course over one school year. This led to small improvements in mathematics and reading (effect sizes of +0.10 for mathematics, and +0.12 for reading), but no change in writing

[1] Text in this paragraph adapted from Alexander (2017).

scores, for the sample as a whole. However, pupils eligible for free school meals saw larger improvements, with effect sizes of +0.20 for mathematics, +0.29 for reading, and +0.17 for writing. The evaluation team for the project suggested that this could be an underestimate as the duration of the project may not have been enough for the full impact to be achieved.

The evaluation reported here represents a much larger study than has been carried out to date, both in terms of sample size and in curriculum coverage. The Dialogic Teaching intervention evaluated in this project aimed to improve the quantity and quality of classroom talk across the curriculum (including, but not exclusive to, mathematics, English, and science) over two school terms. This contrasts with studies mentioned above, such as Gorard, Siddiqui and Huat See (2016) where the intervention took place in one hour per week over one year and focused on discussion of philosophical questions, and Mercer and Sams (2006) where the intervention took place over the course of twelve lessons and focused on particular aspects of the mathematics curriculum.

Evaluation objectives

The research question guiding the impact evaluation was:

- Does the Dialogic Teaching intervention improve pupil attainment in English, mathematics, and science after two terms?

A second question, "Does the Dialogic Teaching intervention improve pupil attainment in English and mathematics after 2 years", will be answered in a forthcoming addendum to this report (expected 2018 publication).

The theory of change for this intervention suggests that changes in teaching practice will lead to increases in the quality of pupil engagement and pupil spoken language observable in the classroom. These changes in pupils' responses to classroom activity are, in turn, predicted to raise levels of attainment in English, mathematics, and science.

With the process evaluation, we aimed to answer two related questions:

- What are the relationships between the training programme, teachers' changing practice, changing classroom interactions, and pupil outcomes?
- Are there differences in the way that the intervention has been implemented by teachers in different schools?

In addition to this process evaluation, the project team undertook separate research assessing teachers' changing practice using analysis of videos. Interim findings are covered in the process evaluation section of this report, and full findings will be published separately.

Ethical review

The evaluation study was approved by the Development and Society faculty ethics committee of Sheffield Hallam University. The ethics submission included those aspects of the study for which the evaluation team was directly responsible, including arrangements for administering post-tests, surveying and interviewing teachers and school leaders, and classroom observations.

Ethical approval for the intervention was secured separately, by the delivery team at the University of York. This covered recruitment of participating schools and delivery of the project. There was co-ordination of the two teams to ensure that all aspects of the project were given consideration with regard to ethical issues arising. A key point of overlap for the two teams concerned consent procedures for the post-test. The delivery team co-ordinated the administration of schools' consent, via headteachers (memorandum of understanding included as Appendix C), and of parental opt-out consent (letter to

parents included as Appendix D), so that parents of children in all participating classrooms were given the opportunity to withdraw their child from the study, including the post-test.

Project team

Project delivery team:

- Professor Robin Alexander: co-director and intervention lead.
- Professor Frank Hardman: co-director and research lead.
- Dr Jan Hardman: discourse analysis lead.
- Dr Taha Rajab: research fellow.
- David Reedy, school liaison officer (pilot stage).
- Mark Longmore: school liaison officer (trial stage).

Evaluation team:

- Professor Tim Jay: principal investigator.
- Ben Willis: project manager, process evaluation.
- Dr Peter Thomas: statistical lead.
- Dr Roberta Taylor: process evaluation, case studies.
- Dr Nick Moore: process evaluation, case studies.
- Professor Cathy Burnett: adviser.
- Professor Guy Merchant: adviser.
- Anna Stevens: research associate.

Trial registration

The trial was registered with ISCRTN as trial number ISRCTN14312500.

Methods

Trial design

This trial employed a three-level (pupils within classes within schools) clustered RCT design. Randomisation[2] was at school level, with half of the schools forming the intervention group and half of the schools forming a control group.

Schools in the control condition were asked to engage in 'business as usual' for the duration of the project. Once the project ended, schools in the control condition were invited to take part in the training sessions and to use the video and audio recording equipment necessary for introduction of a Dialogic Teaching approach in their classrooms. Training and equipment were offered with the proviso that the approach would not be used with pupils in Year 5 in the 2015/2016 academic year (Year 6 in 2016/2017). This was to avoid contamination prior to the analysis of KS2 outcomes from June 2017 national tests.

Participant selection

Schools within Leeds and Bradford local authorities, and Birmingham Education Partnership, were invited to participate in the trial. Recruitment was led by the delivery team (Robin Alexander from CPRT, supported by Mark Longmore from the University of York). Eligible schools were those having at least two Year 5 classes and a high proportion (over 20%) of children eligible for free school meals ('FSM pupils'). The target number of schools was 80: 78 schools were recruited, with two pairs of these acting as federated schools, thus in this document we refer to 76 participating schools in total.

Informed consent was secured at two levels. Headteachers gave informed consent on behalf of their schools (Appendix C), and parents were given the opportunity to withdraw their children from the study (Appendix D).

Outcomes measures

Primary outcomes

The intervention is designed to raise attainment across the curriculum. The three primary outcome measures selected for this trial therefore addressed the three core subjects in the curriculum—English, maths, and science. For the intervention to be considered to have had an effect on attainment across the curriculum, there should be evidence of a positive effect on all three measures.

Pupils were randomised at classroom level to participate in one of the three post-test assessments. All pupils in the study completed one, and only one, assessment. This approach was taken as it reduced the testing burden on pupils and teachers without significantly reducing the statistical power of the analysis.

For each subject area—English, maths, and science—level 10 of the corresponding GL Assessment test, suitable for pupils in Year 5 in the summer term, was used. Each test is standardised for U.K. populations.

- English attainment was measured using the Progress Test in English.[3] This test provides measures of reading accuracy, fluency, and comprehension and was administered as a pencil

[2] Strictly speaking, minimisation (used in this study) is not randomisation as it uses a deterministic process to assign cases to groups. However, 'randomisation' is used throughout for simplicity.
[3] https://www.gl-assessment.co.uk/products/progress-test-in-english-pte/

and paper test in groups. The main body of the test was administered, excluding sections on spelling, punctuation, and grammar.

- Maths attainment was measured using the Progress Test in Maths.[4] This test provides measures of fluency in facts and procedures, fluency in conceptual understanding, and mathematical reasoning and problem solving. The main body of the test was administered, excluding the section on mental arithmetic.

- Science attainment was measured using the Progress Test in Science.[5] This test provides measures of knowledge, and ability to work scientifically in physics, chemistry, and biology. In this case the whole test was administered.

For the English and maths assessments, sections of the tests were excluded for two reasons. The main reason for excluding the spelling, grammar and punctuation, and mental arithmetic sections of the tests was that there was limited theoretical basis for a hypothesis linking attainment in these areas with the Dialogic Teaching intervention. The second reason was that these sections of the tests involved more complex administration on the part of the teacher (including the playing of recordings, and managing timings) which would mean that it would not be possible to administer the three tests concurrently within a classroom.

Participating teachers in all schools provided class lists to the evaluation team. These were used to assign children to each of the three tests: the first child on the list took the mathematics test, the second took English, the third science, the fourth maths, and so on. Packages of tests were distributed to schools with children's names pre-printed on tests to ensure that children took the correct test. Tests were administered by pupils' usual classroom teacher, following instructions provided by the evaluation team. Teachers identified a suitable time to carry out tests within the two-week window beginning 16 May 2016. The evaluation team kept a record of testing dates in all schools. Spot checks were carried out by the evaluation team in a random sample of ten schools to ensure that all instructions were followed. No inconsistencies were identified.

Tests were scored by GL Assessment. Scoring was blinded: assessors had no way of knowing which tests came from intervention schools and which came from control schools. For each measure, the raw, unstandardised, score was used in analyses. It was not possible to use standardised scores as not all sections of each test were completed. Separate analyses were carried out for each of the three primary outcome measures.

Secondary outcomes

Secondary outcomes are Key Stage 2 (KS2) point scores in English and maths. These tests will be taken by pupils in May 2017, approximately two years after the start of the intervention. Data will be obtained from the National Pupil Database when it becomes available (unamended data is expected to be available in September 2017). However, it is understood that the effects of the intervention are likely to be mitigated, and possibly compromised, by (a) change of teacher (from a teacher trained in Dialogic Teaching to one who is not) from Year 5 to Year 6, and (b) a change in pedagogy as pupils approach the KS2 SATs (less extended dialogue, more IRE and text-based teaching).

Analyses of KS2 scores will be included in a future publication, an addendum to this report.

The project team carried out their own evaluation of teacher and pupil talk in intervention and control schools, before and after the intervention. Interim findings from this evaluation are included below as

[4] https://www.gl-assessment.co.uk/products/progress-test-in-maths-ptm/
[5] https://www.gl-assessment.co.uk/products/progress-test-in-science-pts/

part of the process evaluation (the project team have produced an interim report as Alexander with Hardman, 2017).

Sample size

Power analyses for the determination of sample size were carried out using the Optimal Design application (Raudenbush, 2011). A power analysis for this three-level cluster RCT design—with 80 schools, two classes per school, and ten pupils per class—was carried out. We assumed that the Intra-Cluster Correlation Coefficient (ICC) for both class and school levels was 0.1 and that KS1 attainment scores (used as a baseline measure) accounted for 36% of the variance in the primary outcome measure (r = 0.6). The results of this analysis suggest an MDES (minimum detectable effect size) of approximately 0.25 with power of 0.8.

Power calculations were also carried out for this design for the subgroup of FSM pupils. Assuming two FSM pupils per class (~20%), and other assumptions as above, the MDES is approximately 0.42.

A sample size of 80 schools was therefore selected as an appropriate compromise to maximise the chance of observing an effect, should there be one, without compromising the quality of intervention delivery. A large number of additional schools would have been required to lower the MDES further. The project delivery team was able to recruit a total of 78 schools. Two pairs of these were acting as federated schools, thus in this document we refer to 76 participating schools in total.

Randomisation

Minimisation methods were employed in order to achieve balance across intervention and control groups. The MinimPy software package (Saghaei and Saghaei, 2011) was used. Minimisation was conducted by Prof Tim Jay, as principal investigator, and was carried out at school level. Schools were allocated to the treatment or control group using the minimisation method, based on the percentage of FSM pupils within each school, the percentage of pupils using English as an additional language ('EAL pupils'), and school-level KS2 total point score in 2013/2014. For each variable entered into the minimisation procedure, a median split was used to designate schools as either 'high' or 'low' for that measure.

The minimisation produced a group of 38 intervention schools and 38 control schools.

A systematic sampling approach was employed in order to allocate all participating pupils to a test condition. Teachers provided class lists to the evaluation team in alphabetical order by surname. Pupils were then assigned alternately to the maths, English, or science test conditions, so that the first pupil on the list took the maths test, the second took English, the third science, the fourth maths, and so on.

Analysis

A statistical analysis plan was published in advance of this report.[6]

Primary intention-to-treat (ITT) analysis

The primary analysis of the impact of the intervention was a multilevel linear regression model of each primary outcome measure. These models had the primary outcome measure (maths, English, science attainment) as the dependent variable, with the following covariates included:

- school level—membership of the intervention group;

[6] Available at: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/Round_6-_Dialogic_Teaching_SAP.pdf

- pupil participant level—KS1 total point score (from NPD); and
- pupil participant level—FSM eligibility.

Modelling was conducted in stages—a main effects stage followed by an interaction stage. The main effects models included the intervention/control dummy variable along with the KS1 pre-test measure and FSM dummies. The main effects model assumed that any impact of the intervention was consistent across different participant subgroups (for example, FSM and non-FSM participants). To explore whether this was a reasonable assumption, interaction terms were introduced one at a time. The interaction terms included were:

- FSM*intervention (isolating FSM participants who received the intervention). These models explored whether the impact of the programme depended upon FSM status. For example, the programme could have had a greater impact for FSM participants (indicated by a positive coefficient on the interaction term) or it might have a greater impact among non-FSM participants (negative coefficient on the interaction term).
- KS1 point score*intervention (isolating the pre-test scores of intervention group participants). These models explored whether the impact of the programme under evaluation depended upon prior attainment (at KS1). For example, the programme might have had a greater impact among higher attainers (indicated by a positive coefficient on the interaction term) or among lower attainers (negative coefficient on the interaction term). If an interaction term was found to be statistically significant, a subgroup analysis was used to explore this in more depth.

Analysis was conducted using Stata v13.

Missing data

Missing data presents a problem for analysis, whether a pupil is missing a value for an outcome variable (post-test score) or for covariates (for example pre-test score). If outcome data is 'missing at random' given a set of covariates then the analysis has reduced power to detect an effect; if data is 'missing not at random' (for example, differential dropout in the intervention and control groups for unobserved reasons) then omitting these pupils (as in the primary 'completers' analysis) could bias the results. Imputing missing data could improve the robustness of the analysis and examine how sensitive the results are to alternative assumptions.

Seven schools from the intervention group did not return primary outcome measures data. This could potentially have biased the results of the primary analysis. Comparisons between the 31 intervention schools that returned data, and the 38 control schools, were carried out using t-tests for school-level variables including the proportion of pupils eligible for FSM, and 2013/2014 KS2 attainment. This provided a check on balance at post-test.

Secondary outcome analyses

An addendum to the main report will be completed in December 2017, including analysis of KS2 attainment.

At the end of the second year, we will again employ the three-level clustered design, with KS2 point scores for English and maths as outcome variables, KS1 scores as covariate, and experimental group and FSM eligibility as predictors. KS2 attainment tests will be taken by pupils in May 2017, approximately two years after the start of the intervention. Data will be obtained from the NPD when it becomes available (unamended data is expected to be available in September 2017).

For the secondary outcomes, data on all 38 intervention and 38 control schools will be obtained from the NPD.

On-treatment analysis

All schools that completed and returned primary outcome measures participated fully in the intervention. Therefore, analysis of primary outcome measures as described above will in itself constitute an on-treatment analysis.

For the secondary outcome analysis, an on-treatment analysis will be conducted that includes only those 31 schools that participated fully in the intervention for the two terms of the project.

Subgroup analyses

Subgroup analyses of FSM pupils were carried out for the primary analysis, and will be carried out for the secondary analysis. Tests for interaction were conducted to assess whether there were differential effects for FSM pupils relative to other pupils. An analysis of FSM pupils will be carried out whether or not there is a significant interaction as this is EEF's target group.

Subgroup analyses of pupils with low prior attainment will be carried out for both primary and secondary analyses if a significant interaction is found. Tests for interaction will be conducted to assess whether there are differential outcomes for children with low prior attainment relative to other pupils.

Implementation and process evaluation

The evaluation team carried out a survey and interviews with headteachers, mentors, and teachers in intervention schools, plus detailed case studies, including classroom observation, of a sample of three intervention schools. In addition to these sets of data, the project delivery team also collected records of attendance at training events and evidence of schools' completion of mentor meeting cycles.

The survey for headteachers, mentors, and teachers in intervention schools addressed research questions relating to the effectiveness of the training provided by the project delivery team, changes in teaching practice, and the perceived effects on classroom interactions, pupil engagement, attitudes, and attainment. Surveys emphasised aspects specific to each role, so, for example, the survey for teachers emphasised changes in pedagogy and pupil behaviour, the mentor survey included questions concerning the conduct of mentor meetings, and the survey for headteachers emphasised whole-school aspects of the intervention.

All classroom teachers, mentors, and headteachers in intervention schools were asked to complete these surveys. Completed surveys were returned by 12 headteachers, 16 mentors, and 24 teachers. Telephone interviews were carried out with six teachers, eight mentors, and three headteachers from a sample of eight intervention schools. These were designed to further probe the topics covered in the surveys. The sample for the interviews was selected at random from the set of schools that did not return completed surveys.

Case study methodology

This aspect of the process evaluation was designed to examine how the intervention was implemented in classrooms. Three schools were selected as case study schools, in collaboration with the project delivery team's school liaison officer. Schools were selected as representative of those in which, from the perspective of the delivery team, the intervention had been working well. The selected schools represented inner city and suburban populations, and there was a diversity of ethnic backgrounds amongst the students.

- School A

  This primary school is in an inner-city area of a medium-sized city. It is much larger than average. School A is characterised by high proportions of FSM and EAL pupils, a wide range of heritage languages, and a higher-than-average number of students joining the school during the year. The

mentor at this school was the deputy head, and three experienced teachers were recruited to the project.

- School B

This primary school is located on the edge of a small, post-industrial city. It is larger than average and has experienced a lot of recent growth. The proportion of pupils at this school eligible for pupil premium and for school action plus, as well as those with an SEN statement, is higher than most schools in the U.K. The majority of pupils come from a white British background. The deputy head took the role of mentor in this school and two experienced teachers were recruited to the project.

- School C

This larger-than-average primary school is located in an inner-city area of one the U.K.'s largest cities. More than double the national average number of pupils at this school qualify for pupil premium support. Pupils at this school speak more than 30 different languages, and one in four arrives at school with little or no English. In school C, the mentor role was taken by a member of the senior management team at the school. In this school the project recruited three teachers, including one NQT.

An observation schedule was designed to focus on questions and dialogue in the classroom. This was not intended as a measure of the extent to which Dialogic Teaching had been implemented in classrooms, but as a framework to guide observers in their attention to features of Dialogic Teaching. An interview schedule designed to elicit teacher and mentor perceptions of the programme, and their views on its efficacy, was also prepared. Observation schedules were completed during the classes, and interviews and observations were either recorded and later transcribed, or notes were taken, depending on the school's and teacher's preferences. Documentary evidence, such as revised Schemes of Work, was also collected where available.

The data set was analysed using six questions based upon the project aims as set out in the 'Handbook for the Development Phase Schools' (Alexander and Hardman, 2014, p. 2).

- What are the effects of the strategy for maximising the quality and educational impact of classroom talk using a Dialogic Teaching approach?
- How has an environment which fosters pupil attention and interest been promoted in the case study schools?
- What does the data-set tell us about talk which meets and goes beyond the requirements of the National Curriculum and what examples are there of pupils using talk to reason, explain, justify, argue, speculate, evaluate, and in other ways think for themselves?
- What evidence is there of the Dialogic Teaching approach being used across the curriculum in the case-study schools?
- What are the teachers' and mentors' perceptions of the ability of the programme to raise standards in literacy, oracy, and higher-order thinking?
- What are the mentors' views of the effectiveness of this model of professional development, its sustainability, and its potential to be scaled up?

Development team process evaluation

The project development team (the Cambridge Primary Review Trust with the University of York) carried out an in-house evaluation, which had two strands: an interview programme undertaken in intervention schools only, and a comparative analysis of videotaped lessons in both intervention and control schools. This work is separate from the evaluation carried out by the Sheffield Hallam University (SHU) team and has not been scrutinised or validated by SHU as independent evaluators. It will be published separately in full after this evaluation report. An interim analysis of lesson videos is reported under 'Process Evaluation' below. The paragraphs below are taken from the development team's interim report (Alexander with Hardman, 2017).

"In order to assess the pedagogic impact of the intervention, lessons were video recorded in a sample of both the intervention and control groups. Lessons were recorded twice, so as to assess development and progress over time. Video recordings of a sample of English, maths, and science lessons were made (1) at the beginning of the trial (week beginning 21 September 2015) to provide a baseline and (2) towards the end of the trial (fortnight beginning 22 February 2016). Fifteen teachers from the intervention group and 11 from the control group agreed to be video-recorded. The intervention group teachers were self-selected in response to our request for volunteers at the July 2015 induction session. The control group teachers were then selected on the basis of school-to-school matching. Each teacher was recorded twice—in phase 1 and again in phase 2[7]—yielding a theoretical total of 156 lessons (two English, two maths, and two science in each case). In fact, because not all of the designated teachers taught science, the total number of lessons recorded was 134 (67 in each phase). The resulting recordings were subjected to both quantitative and qualitative analysis (quantitative analysis only reported in this report).

For the quantitative analysis, some of the key verbal indicators of typical classroom talk, both dialogic and traditional, became the basis for a coding system that was piloted in a previous study before being finalised and applied to the trial stage video data. Coders were trained and checked to maximise coding consistency. The coding system for these was uploaded into the Observer XT 12.5 software (Zimmerman *et al*., 2009) in order to generate quantitative data from the coded acts and exchanges. These were then statistically analysed using SPSS. The analysis was undertaken twice for the purpose of cross-validation, first internally at the University of York, then externally by Kirkdale Geometrics."

Costs

Cost information was provided by the development team for training sessions, equipment, and other resources based on a school with two Year 5 classes (the average in this trial).

Questions were also included in the survey carried out by the evaluation team in order to collect data regarding the amount of time required for mentor meetings and additional planning time for mentors and teachers.

Per-pupil costs were calculated with the assumption that there were 30 pupils in each class (and so 60 per school).

Timeline

Table 2 below summarises the main activities undertaken by the evaluation team and project delivery team relating to this trial.

[7] See point 5 on page 7.

Table 2: Timeline

Date	Activity
February–May 2015	**Recruitment of participants** Project delivery team co-ordinated recruitment, via local authority education leads.
June 2015	**Minimisation** Evaluation team minimised recruited schools to two groups. **Induction training session** Headteachers, mentors, and teachers in the intervention group attended induction training sessions led by the project delivery team. Teachers received training materials and audio/video recording equipment.
September 2015	**Mentor training session** Headteachers and mentors attended a training session led by the project delivery team. The project delivery team also arranged 'catch-up' induction sessions for any schools that were not able to attend the June session.
September 2015–April 2016 (autumn and spring terms)	**Implementation** Teachers in intervention schools using a Dialogic Teaching approach in their classrooms. Mentor meetings occurring fortnightly. Monitoring and support provided by the project delivery team.
December 2015	**Mid-intervention plenary** Project delivery team led plenary session for intervention schools.
January 2016	**Survey** Evaluation team distributed surveys for headteachers, mentors, and teachers in intervention schools.
February–March 2016	**Interviews** Evaluation team conducted telephone interviews with a sample of headteachers, mentors, and teachers from intervention schools. **Case study visits** Evaluation team visited three schools for collection of data relating to the case studies.
May 2016	**Plenary sessions** Project delivery team led plenary sessions for intervention schools following the end of the intervention period. **Post-test** Testing of maths, English, and science attainment was carried out in intervention and control schools.

Impact evaluation

Participants

Figure 1 outlines the flow of participants during the course of the project. Six schools withdrew from the intervention group at a very early stage, before the intervention had begun, and a further school from this group completed the intervention but did not complete post-tests. Complete datasets were not available for 265 pupils in the intervention group and 386 pupils in the control group, either because post-tests were not completed (due to absence or opt-out), or because pupils for whom post-tests were returned by schools could not be matched with data from the NPD regarding KS1 attainment and FSM eligibility. In total, complete data was available for 1,832 pupils in intervention schools and 2,080 pupils in control schools.

Table 3 shows the minimum detectable effect size (MDES) at various stages in the development of the trial. It also shows that estimates of the correlation between pre-test and post-tests, and of intra-cluster correlations, were reasonable and not substantially different from observed values.

Figure 1: Participant flow diagram

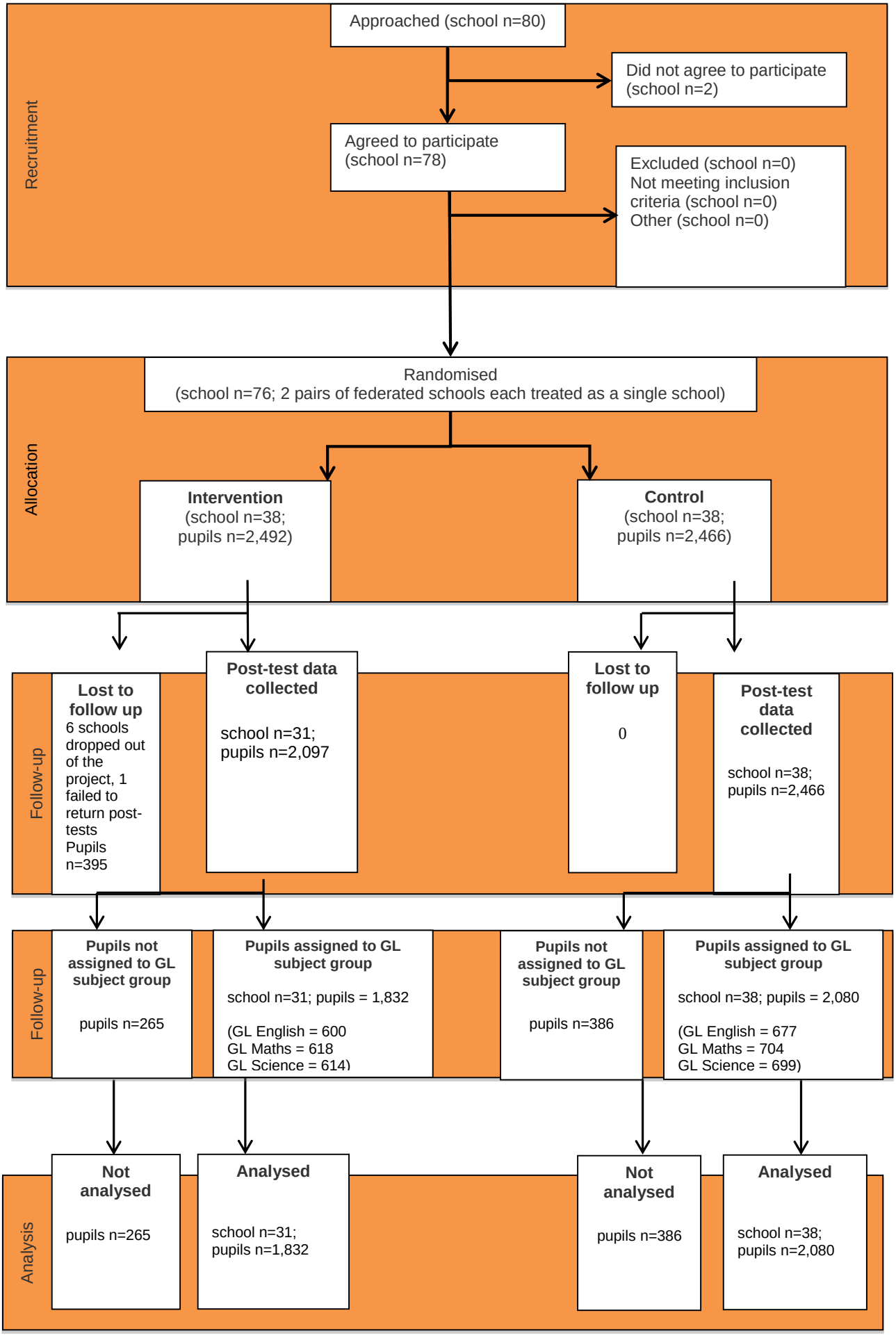


Table 3: Minimum detectable effect size at different stages

Stage	N [schools] (n=intervention; n=control)	Correlation between pre-test (+other covariates) and post-test	ICC	Blocking/ stratification or pair matching	Power	Alpha	Minimum detectable effect size (MDES)
Protocol	80 (40; 40)	0.60	0.10	minimisation	80%	0.05	0.25
Randomisation	76 (38; 38)	0.54	0.10	minimisation on FSM, EAL, and KS2	80%	0.05	0.26
Analysis (i.e. available pre- and post-test)	69 (31; 38)	0.54	0.10	minimisation on FSM, EAL, and KS2	80%	0.05	0.28

The MDES was calculated using Optimal Design software on the basis of a three-level cluster random design (schools, classes, pupils) with treatment at level three. Calculations were based on two classes per school and ten pupils per class. Inter-Class Correlations are on the basis of schools and classes.

Baseline characteristics

Table 4 summarises the school characteristics of the sample at the analysis stage. The obvious notable imbalance is that there are 31 intervention schools at the analysis stage compared to 38 at randomisation. The control group remains unchanged with 38 schools remaining (see Appendix E, Table 1 for further details). That said, there remain similar school characteristics between intervention and control schools at the analysis stage. There are similar distributions of Ofsted ratings between groups (and schools that dropped out of the intervention group had a similar distribution of Ofsted rating too—see Appendix E, Table 1). The table indicates that the percentages relating to pupils' gender, SEN support status, and FSM eligibility are identical between the intervention and control groups. The one notable difference concerns EAL pupils: here, at school level, 53% of pupils in the intervention schools are categorised as EAL pupils compared to 47% in control schools. Across both groups, there is a high proportion of EAL pupils: this could have some implications for external validity.

Table 4: School characteristics at analysis (intervention 31 schools, control 38 schools)

Variable School-level (categorical)	Intervention group n/N (missing) / Percentage		Control group n/N (missing) / Percentage	
School type				
LEA Maintained	24/31	77%	30/38	79%
Academies	7/31	23%	8/38	21%
Ofsted rating	30/31 (1)*	97%	37/38 (1)*	97%
Outstanding	4/30	13%	5/37	14%
Good	19/30	63%	25/37	67%
Requires improvement	6/30	20%	7/37	19%
Inadequate	1/30	3%	0/37	0

EverFSM	31/31	35%	37/38	35%
EAL	31/31	53%	37/38	47%
Pupils with SEN Support	31/31	16%	38/38	16%
Gender				
Female	31/31	49%	38/38	49%
Male	31/31	51%	38/38	51%

***** Ofsted rating not available for 2 of the 76 schools. This is generally due to changes in status/name of a school. For example, when a school converts to Academy status, an Ofsted rating will not be available until it has been inspected after conversion.

Table 5 summarises the pupil characteristics of the sample at analysis stage. The first point to note is that there are more pupils in the control group (53%) than the intervention group (47%), representing an imbalance from the original randomisation when pupils were split virtually 50/50 between intervention and control group (see Appendix E, Table 2). At randomisation there were 2,492 pupils in the intervention group and 2,466 in the control group. Of the former, 1,256 were FSM pupils (50.4%), 1,209 female (48.5%), and 1,283 male (51.5%). Of the 2,466 control group pupils, 1,221 were FSM pupils (49.5%), 1,227 female (49.8%), and 1,239 male (50.2%).

By analysis, the overall number of pupils in the intervention group had reduced from 2,492 to 1,832 (meaning 660 missing cases). Breaking this down by pupil category, FSM numbers reduced from 1,256 to 917 (339 FSM pupils not analysed), females from 1,209 to 907 (302 not analysed), and males from 1,283 to 925 (358 not analysed).

The overall number of pupils in the control group had reduced from 2,466 to 2,080 (386 missing cases). Breaking this down by pupil category, FSM pupils went down from 1,221 pupils to 1,021 (200 not analysed), females from 1,227 to 1,051 (176 not analysed), and males from 1,239 to 1,029 in the final analysis (210 not analysed).

In terms of pupils in the final analysis, the intervention group had a slightly higher proportion of boys than girls, whereas the control group had a slightly higher proportion of girls than boys. At 50.1% (intervention) and 49.1% (control), the proportion of pupils ever eligible for FSM was slightly higher in the intervention group. The assigning of pupils into GL Assessment tests is very similar between groups, with GL maths being the subject most pupils were assigned to for both groups, followed by science, and finally English—the only subject where pupils assigned to outcome measure falls below 33% for both intervention and control groups. In terms of the NPD KS1 mean scores, the control group is slightly higher than the intervention group both in terms of all pupils and when we account for 'Ever FSM' pupils only.

Table 5: Pupil characteristics at analysis (intervention 31 schools, control 38 schools)

Pupil-level (categorical)	Intervention		Control	
	n/N (missing)	Percentage	n/N (missing)	Percentage
Number of Y5 pupils	1,832/3,912 (660)	47	2,080/3,912 (386)	53
Eligible FSM	917/1,832	50.1	1,021/2,080	49.1
Gender				
Female	907/1,832	49.5	1,051/2,080	50.5
Male	925/1,832	50.5	1,029/2,080	49.5

Pupils assigned to GL subject				
GL English	600/1,832	32.8	677/2,080	32.6
GL Maths	618/1,832	33.7	704/2,080	33.8
GL Science	614/1,832	33.5	699/2,080	33.6
Pupil-level (continuous)	**n (missing)**	**Mean**	**n (missing)**	**Mean**
NPD KS1 score	1,725 (107)	14.838	1,992 (88)	15.017
NPD KS1 score (Ever FSM)	882 (35)	14.400	989 (32)	14.572

Table 6 examines the pupil-level baseline balance for those pupils included in each of the three outcome measures in order to explore potential bias introduced at assignment to outcome measure. In Table 5, we see that overall percentage of Ever FSM pupils in the intervention group is 50.1%. Exploring across outcome measures we can see there is only very minor variation across subjects (a range of 49.7% to 50.5%). The proportion of Ever FSM pupils in the control group is 49.1%. Exploring across outcome measures we can see there is slightly more variation across subjects (a range of 47.8% to 50.6%).

Similarly, for the intervention group the KS1 score for all pupils is 14.838. Exploring across outcome measures we can see there is only very minor variation across subjects (a range of 14.817 to 14.826). We also see this minor variation repeated for Ever FSM pupils. For the control group, the KS1 score for all pupils is 15.017. Exploring across outcome measures we can see there is minor variation across subjects, but again this is slightly more pronounced for the control group compared to the intervention group (a range of 14.894 to 15.090). We also see this minor variation repeated for Ever FSM pupils.

Table 6: Pupil characteristics by GL Assessment outcome measure

Pupil-level by outcome Categorical (GL English)	**n/N**	**Percentage**	**n/N**	**Percentage**
EverFSM	303/600	50.5	331/677	48.9
Female	318/600	53.0	348/677	51.4
Male	282/600	47.0	329/677	48.6
Pupil-level by outcome Categorical (GL Maths)	**n/N**	**Percentage**	**n/N**	**Percentage**
EverFSM	307/618	49.7	356/704	50.6
Female	292/618	47.2	360/704	51.1
Male	326/618	52.8	344/704	48.9
Pupil-level by outcome Categorical (GL Science)	**n/N**	**Percentage**	**n/N**	**Percentage**
EverFSM	307/614	50.0	334/699	47.8
Female	297/614	48.4	343/699	49.1
Male	317/614	51.6	356/699	50.9
Pupil-level continuous (GL English)	**n (missing)**	**Mean**	**n (missing)**	**Mean**
NPD KS1 score	569 (31)	14.817	648 (29)	15.066
NPD KS1 score (Ever FSM)	294 (9)	14.337	320 (11)	14.759
Pupil-level continuous (GL Maths)	**n (missing)**	**Mean**	**n (missing)**	**Mean**
NPD KS1 score	578 (40)	14.876	681 (23)	15.090
NPD KS1 score (Ever FSM)	290 (17)	14.500	347 (9)	14.473
Pupil-level continuous (GL Science)	**n (missing)**	**Mean**	**n (missing)**	**Mean**

NPD KS1 score	578 (36)	14.821	663 (36)	14.894
NPD KS1 score (Ever FSM)	298 (9)	14.364	322 (12)	14.484

Outcomes and analysis

Table 7: Primary analysis, controlling for KS1 prior attainment, pupil Ever FSM, and school-level variables—FSM eligibility, EAL, and KS2

	Intervention group		Control group				
Outcome	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**n in model (intervention; control)**	**Effect (95% CI)**	**p-value**
Post-test GL Progress Test English	600	13.76 (6.178)	677	13.16 (6.057)	1,198 (569; 629)	+0.15 (-0.001, 0.299)	0.05
Post-test GL Progress Test Maths	618	21.25 (10.843)	704	20.98 (10.347)	1,239 (577; 662)	+0.09 (-0.039, 0.199)	0.19
Post-test GL Progress Test Science	614	26.67 (8.227)	699	26.29 (8.240)	1,223 (578; 645)	+0.12 (0.007, 0.226)	0.04

The primary analysis of the impact of the intervention was a multilevel linear regression model of each primary outcome measure. These models had the primary outcome measure (GL maths, GL English, and GL science attainment) as the dependent variable, with the following covariates included:

- school level (variables included in minimisation)—membership of the intervention group, percentage of FSM pupils within each school, percentage of EAL pupils, school-level average KS2 total point score 2013/2014;
- pupil participant level—KS1 total point score (from NPD); and
- pupil participant level—FSM status.

The main effects model is used for the main results. This was constructed in two steps:

1. pupil participant variables—intervention/control dummy variable and KS1 attainment; and
2. as 1, but also including FSM status (pupil level) and the school-level variables included in minimisation (% FSM; % EAL, and school-level KS2 attainment).

The second step represents the models that will be used for the main results for the three primary outcomes for this trial. The first step is included for sensitivity and so that the findings can be comparable across different trials. Covariate tables are reported in Appendix F.

Effect sizes for a cluster randomised trial using total variance can be calculated as:

$$ES = \frac{(\bar{Y}_T - \bar{Y}_C)_{adjusted}}{\sqrt{\sigma_S^2 + \sigma_{error}^2}}$$

Where (ȲT−ȲC) adjusted denotes ANCOVA mean difference between intervention groups adjusting for NPD KS1 pre-test score and other pupil background variables (Ever FSM) and minimisation variables (when necessary).

Main effects stage

When examining the main effects models—controlling for prior attainment, pupil FSM and school-level FSM eligibility, EAL and KS2 (reported in Table 7)—there is a statistically significant result for science. The post-test marks for the intervention group in GL science were 0.38 units higher than the control. It is unlikely that this result has happened by chance (p = 0.04). There was a modest effect size of +0.12. It should be noted that this was lower than the minimum detectable effect size (MDES) calculated in the power analysis. However, it is important to note that the power calculations in Table 3 are based on the correlation between KS1 prior attainment pre-test and post-test only, not all variance explained by covariates in the primary analysis model. Given that statistically significant results reported in the main effects stage occurred when we introduced the school-level minimisation covariates this is likely to account for why we find some statistically significant results in the primary analysis but have a higher MDES in the power calculations.

There is a cautiously positive result for English. The post-test marks for the intervention group in GL English were 0.59 units higher than the control. The results are very close to being called statistically significant but we cannot discount the possibility that this result has happened by chance (p = 0.051). There was a modest effect size of +0.15. Again, this was lower than the minimum detectable effect size (MDES) calculated in the power analysis.

The post-test marks for the intervention group in GL maths were 0.27 units higher than the control. We cannot discount the possibility that this result has happened by chance (p = 0.19). There was a low effect size of +0.09.

Sensitivity analysis

When examining the main effects models considering intervention and control dummy, and controlling for KS1 prior attainment only, modest effect sizes were shown for GL English (+0.14), GL maths (+0.08), and GL science (+0.08) and it is not possible to rule out chance as an explanation for the effect observed. Thus, for the three main outcomes, there were no statistically significant results when observing intervention and prior attainment only.

Turning to sub-group analysis for free school meals pupils, when observing only intervention and prior attainment, again there were only modest effect sizes shown for GL English (+0.12) and GL science (+0.10) and it is not possible to rule out chance as an explanation for the effect observed. However, the effect for GL maths (1.39 units higher than the control, effect size +0.16) can be reported as statistically significant (p = 0.02).

Sub-group analysis

The analyses performed for the full data were repeated for the subgroup of FSM pupils (using EverFSM).

For FSM pupils, modest effect sizes were shown for GL English (+0.12) and GL science (+0.11) and it is not possible to rule out chance as an explanation for the effect observed. However, the effect for GL maths (1.39 units higher than the control, effect size +0.16) can be reported as statistically significant (p = 0.03). Again, this was lower than the MDES calculated in the power analysis. Table 8 summarises the outcomes for FSM pupils. In Table 8 missing data is not reported for separate GL Assessment tests because it does not exist. The missing data is as follows:

Beginning with the intervention group, from randomisation to analysis there are 339 missing cases. If we look at follow-up stage (with seven intervention schools no longer part of the trial) then 1,036 Ever FSM intervention pupils remained in the trial, of which 917 were assigned to GL Assessment tests, meaning 119 pupils who remained at follow-up were not part of the final analysis.

Turning to the control group, from randomisation to analysis there are 200 missing cases. Between randomisation and follow-up no cases were missing, meaning 1,221 Ever FSM control pupils remained in the trial. Of these, 1,021 were assigned to GL Assessment tests, meaning 200 Ever FSM control pupils who remained at follow up were not part of the final analysis.

Table 8: Free school meal eligible (Ever FSM) pupils, controlling for KS1 prior attainment and school-level variables—FSM eligibility, EAL, and KS2

	Raw means				Effect size		
	Intervention group		Control group				
Outcome	n	Mean (SD)	n	Mean (SD)	n in model (intervention; control)	Effect (95% CI)	p-value
Post-test (GL Progress Test English)	303	12.87 (6.134)	331	12.63 (6.056)	603 (294; 309)	+0.12 (-0.065, 0.299)	0.21
Post-test (GL Progress Test Maths)	307	19.99 (10.474)	356	18.60 (9.393)	627 (290; 337)	+0.16 (0.014, 0.298)	0.03
Post-test (GL Progress Test Science)	307	25.66 (8.077)	334	25.03 (7.995)	613 (298; 315)	+0.11 (-0.038, 0.260)	0.14

On-treatment

The analysis of the primary outcomes has revealed some statistically significant and potentially promising results. At analysis stage, intervention group pupils uniformly had lower NPD KS1 scores than control group pupils in terms of overall results, by GL subject allocation and by EverFSM. (The only intervention subgroup to have a higher KS1 score than its control subgroup counterpart was for GL Maths EverFSM.) At post-test across all GL Assessment tests, the intervention group attained higher mean average scores than the control group across primary outcomes allocation and subgroup. Given this, exploring the Key Stage 2 results will be of particular importance to further explore the relative strengths of the intervention compared to pupil background variables and school-level factors.

The findings reported here constitute an 'on-treatment' analysis, as data was only available for pupils in those schools in the intervention group that fully participated in the intervention (data was available for all pupils in control schools).

Missing data

It is important to reiterate that the positive findings identified in this project must be treated with caution due to the incidence of missing data. At school level, 7 of the 38 randomised intervention schools did not complete the post-test measure GL Assessment tests (see page 16). Due to the fact that all schools dropping out or not completing were intervention schools, this created an imbalance at follow-up and analysis stage. (Although imputation was considered in the SAP, given the large issue of whole schools dropping out it was not considered feasible in this trial.)

At randomisation stage there were 2,492 pupils (50.3%) randomised into the intervention group and 2,466 pupils (49.7%) into the control group. At the follow-up stage, there were 2,097 pupils in the intervention group (46%), while 2,466 remained in the control group (54%). Following assignment to GL

Assessment tests (the outcome measure), the final numbers at analysis stage were 1,832 intervention pupils (47%) and 2,080 control pupils (53%).

In terms of baseline characteristics, there was already an imbalance at randomisation for the school-level variable 'EAL pupils' between intervention and control schools (52% of pupils at intervention schools were EAL pupils, compared to 47%). Following whole school drop-outs, this increased to 53% for the intervention group. In terms of NPD KS1 mean scores, these increased for intervention pupils from 14.65 at randomisation to 14.84 at analysis; for control pupils this increased from 14.82 at randomisation to 15.02 at analysis. In terms of NPD KS1 mean scores for Ever FSM pupils, these increased for intervention pupils from 14.20 at randomisation to 14.40 at analysis; for control pupils this increased from 14.33 at randomisation to 14.57 at analysis. From an initial 4,958 pupils randomised to intervention and control groups, 3,912 completed post-tests, representing 79% of the original pupils.

Cost

This estimate of cost per school is based on the assumption that two Year 5 teachers, one mentor, and one headteacher from each school would participate in the training. This was the case for the majority of schools in the Dialogic Teaching trial.

For the trial, each teacher attended three days of training, mentors attended four days of training, and the headteacher attended for one day (a total of 11 participant/days of training, assuming two teachers in a school). In addition, each school was visited three times during the course of the two terms of implementation of the mentoring cycles. Schools were also provided with video recording equipment and three additional days of supply cover to facilitate the recording of lessons and to act as a stimulus for critical reflection and discussion during the mentoring sessions.

Financial costs

Training venue/catering costs (11 participants @ £150)	£1650
School visits (3 visits @ £250 per visit)	£750
Training handbooks (5 handbooks @ £20 each)	£100
Video recording equipment (camera, tripod, voice recorder)	£400
Workshop training costs	£200
Total	**£3,100**

At £3,100 per school per year, this gives a cost of approximately £52 per pupil per year, assuming 60 pupils per school.

It is not possible to calculate an accurate cost over three years as it is not clear what form the intervention would take over this period. If the intervention was repeated within a single year group (Year 5 in the current intervention), then costs would depend on whether the teachers in that year group changed between years, which would mean that further training would be required. It is likely that a Dialogic Teaching approach would be implemented at a whole-school level, rather than in a particular year group, and so different assumptions regarding costs would be required. In order to provide a cost per pupil per year over three years, we have therefore used the-per pupil per year cost of £52.

Teaching cover requirement

In addition to the 11 days' cover required for teachers, mentor, and headteacher to attend training sessions, survey data (from 12 headteachers) showed that teachers and mentors each required an average of 24 hours per term (72 hours per year) for mentor meetings and additional preparation and planning time. In total, this represents 37 teacher-days per year that could require cover. However, as we note in the process evaluation findings below, schools managed such meetings in different ways with some making use of existing PPA time and others making use of cover. Cover required in practice was therefore closer to 17 days per year (11 days for training, plus two days per teacher and mentor for meetings and planning).

Process evaluation

This section reports the findings of the process evaluation carried out by the evaluation team, comprising survey, interviews, and case studies. The survey and interviews focused primarily on the implementation of the intervention, including findings pertaining to fidelity, and on participants' perception of outcomes of the intervention. Surveys were completed by 12 headteachers, 16 mentors, and 24 teachers. Interviews were carried out with a further three headteachers, eight mentors, and six teachers. The case studies focused on changes in classroom practice, including teacher and pupil behaviours as a result of participation in the project.

Under 'Outcomes' below we also include an account of analysis carried out by the development team of video observations of teacher and pupil talk. This work was carried out by CPRT and the University of York and has not been scrutinised or validated by the SHU independent evaluation team. Videos of lessons were recorded at the beginning (September 2015) and towards the end (February 2016) of the trial, and analysis compares teacher and pupil talk observed at these two points both over time and between the intervention and control groups.

Throughout this section, respondents are identified by region (A or B), school number, and by their role (headteacher, mentor, or teacher). This is to communicate the range of sources of opinions given, while maintaining participants' anonymity. Unless indicated otherwise, quotations included in this section are broadly representative of participants' responses. Data from the survey and from the interviews agreed in all areas, and so findings and respondents' comments from the two sources are merged below.

A more detailed account of case study findings are included in Appendix G.

Implementation

What are the necessary conditions for success of the intervention?
Senior leadership approval and management of staff time (including arrangements for teaching cover)

The most crucial condition identified was for the school and participants to be able to dedicate sufficient time to the project. Senior leadership buy-in to the intervention presented itself in both direct and indirect ways. For the most part, school leaders, while acknowledging it was difficult to protect precious staff time, made a considered judgement that doing so was a valuable and necessary requirement for being involved in the programme and did all they could to facilitate this, most usefully in the form of providing dedicated teacher cover arrangements.

> *'Head has been very good with regards to non-contact time for us to meet as well. We've generally met to plan and review sessions and to look at recordings every two weeks, so within each cycle, certainly within each cycle'* (mentor, RegB-S5).

There were schools where paid cover was not implemented. For example, one school (RegA-S1) revealed that despite the 'offer of the money to get supply' (mentor), the NQT status of one teacher and the pre-existing training commitments of the other meant that the school was unwilling to sanction additional time out of the classroom 'because the classes need that teacher to be teaching'. In such circumstances, this meant having to find time outside of timetabled teaching to fit in key project requirements such as mentor meetings. Nevertheless, interview and survey data (plus attendance records at events) show that teachers, mentors, and headteachers attended the required face-to-face sessions (induction, training, and plenary sessions) run by the delivery team and to engage in the fortnightly mentoring cycles.

Mentor meetings and teacher commitment

The mentor meeting was routinely identified by both mentors and teachers as being a particularly valued feature of the programme. The role of the mentor was perceived to be of central importance:

> *'I think teachers are so busy with planning, marking, assessment, all of the other things that go on in a school, and to be able to have that time with my mentor, who's the Deputy Head, and just sit down and focus on a certain aspect of my teaching and have her feedback and have a conversation about it and be able to discuss where to go next, has just been really, really helpful, just having that time'* (teacher, RegA-S7).

The regularity and quality of these mentor meetings were linked with senior leadership backing for the project and their willingness to protect staff time for it (as discussed above). Additionally, respondents cited the personal commitment of staff as a critical determinant of the effectiveness of implementation. Survey findings showed that the intervention required a substantial commitment from teachers and mentors in terms of time, with heads estimating that teachers and mentors had spent an additional 24 hours per person per term outside of lesson time on project activities, including mentor meetings. In survey responses, all mentors mentioned time as a major challenge to the implementation of the intervention. It was not easy for all schools to arrange cover for teachers and mentor (although some mentors did not have classroom teaching responsibility), and some schools were also coping with other staffing difficulties.

Project resources and equipment

There was consensus about the quality of all materials provided, although there was some disparity in opinion with regard to the volume and appropriateness of certain materials provided by the delivery team. At one end of the continuum there were individuals claiming that 'everything' that had been provided had been 'useful in one way or another' through to others that seemed slightly overwhelmed, particularly in relation to the more 'academic' content.

The project handbook and laminated 'Dialogic Teaching Repertoires' sheet that included nine key 'talk moves' were regularly singled out as being the most useful and consistently used resources. Engagement with these resources most readily enabled participants to access and engage with the programme in the context of time-pressured school working environments.

> *'I think the handbook was invaluable. That really laid things out'* (mentor, RegA-S2).

> *'The talk-moves and things like that on this teaching repertoire sheet were the most useful, because it was just quick to get down to what we needed to do really'* (mentor, RegB-S4).

Finally, the use of video recorded classroom activity was identified as a particularly powerful feature of the programme, which underpinned the mentor meetings.

> *'I think the tool of using the videos, so all that recording equipment; I think that's been the most valuable aspect of the whole project. You could watch the lesson together and pull things out of that and it's been extremely valuable in making the teachers more reflective'* (mentor, RegA-S2).

Are there any barriers to delivery being experienced?

Time management

Across all schools, mentors and teachers found it difficult on some occasions to find the time to participate in mentor meetings. In some settings, diminished staff capacity resulting from absences

exacerbated this problem. Some teachers also referred to changing timetabling of PPA (planning, preparation, and assessment) time through the period of the intervention, which made arranging mentor meetings more challenging:

> *'There was a period where it was quite stable and we could meet regularly, once a fortnight at the end of every cycle, but then timetables change and it's tricky to find time when we're all available'* (mentor, RegB-S4).

The frequency of the fortnightly learning cycles sometimes presented difficulties, as there were times when teaching of core curriculum subjects was reduced in order to focus on other aspects of pupils' education. One mentor gave the example of sex and relationships education having to take priority which meant that it was not possible to complete project recording and mentor meetings during that period.

> *'I think one of the biggest challenges we've had is trying to fit it in. For me personally going from one cycle to another every two weeks is really hard. We've got sex and relationships education going on for the next three weeks, this week and the next two weeks in Year 5. We cannot fit in the project recordings or the planning and review sessions'* (mentor, RegB-S5).

Some mentors adapted the timings and content of the intervention to create more flexibility. For example, one mentor reported 'adapting' and 'filtering' parts of the programme to make engagement more manageable for teachers:

> *'In sessions where it said video three sessions, I've said to staff just video one and we're going to focus on one'* (mentor, RegB-S3).

Breadth and complexity of information received

Participants were universally positive about the quality and standard of professional development received from the delivery team, both in terms of their direct input and wider resources. There was no suggestion that there was anything significant missing from the delivery provided. However, teachers made reference to there being too much content for participants to fully engage within the context of their busy working lives.

> *'I don't think there is any [additional professional development that would have been helpful]. I think we've had all the help we need. Again, it's just fitting it in to the time constraints of the school'* (mentor, RegA-S1).

For most schools, having a wide range of content and resources to draw upon was seen as positive, and individuals accepted they would not necessarily use all the available resources and features of the programme to the same extent. However, for a minority of participants, the scale of the programme, combined with the breadth of unfamiliar terminology used, was quite intimidating. In the example provided below, a mentor explains how their school's focus was based on decisions about which aspects they found more accessible:

> *'I think it was overwhelming, to be honest. I think the problem is there are so many different terms and words that they're [delivery team] aware of and use all the time, and to us they are all completely new … So I find that a bit daunting, but I'm afraid I'm quite sensible and practical and say, well, do you know what; we'll take what we can from this. We'll take the good things and we'll do what we can do that helps the children, but the rest, you know ...'* (mentor, RegA-S1).

Conflicting agendas and school priorities

The intervention generally aligned with schools' existing priorities. However, there were instances where the Dialogic Teaching programme priorities operated in tension with other school priorities or circumstances. For example:

> *'We're a school with a standards agenda in that we have to be very mindful of that. We've also got a new curriculum in place which we're implementing this year for the first time properly and we've got a new assessment process we're implementing this year for the first time properly. There are a lot of other factors to think about, so the quality of dialogue hasn't necessarily been the priority'* (mentor, RegB-S3).

In a separate example, a mentor reported that despite an underlying belief that the Dialogic Teaching project was important and benefitted pupils in numerous ways, it remained just 'one part of everything' (mentor, RegA-S1). School 1 in Region A had recently embarked on project to introduce a mastery approach in maths, requiring a shorter initial 40-minute teaching session each day followed by a second session later in the day. The mentor reported that this altered format did not lend itself particularly well to a Dialogic Teaching approach and this caused staff to experience some conflict:

> *'So if you've only got 40 minutes in your maths lesson, you're not going to spend 20 minutes doing a beautiful little dialogic discussion'* (mentor, RegA-S1).

In some schools, issues were noted regarding perceived requirements to demonstrate progress in lessons for senior staff and for Ofsted. In some cases this also meant that teachers could be somewhat suspicious of the mentor role and, initially at least, uncomfortable with the observation of footage from their lessons. This was despite reassurances from both mentors and the project development team that the mentor meetings should be unrelated to performance management.

> *'You have to build up that trust and relationship and say actually this is focused for the project and we're looking at talk and engagement and interaction with the pupils and the intervention that you're making, not your teaching as a whole. I'm not here to judge you, you know. You have to break down those barriers and that's a challenge'* (mentor, RegB-S5).

The mentor quoted above also reported that teachers were initially uneasy about a perceived lack of evidence of improvements in children's learning associated with Dialogic Teaching and were instead more concerned about generating sufficient written evidence in books to demonstrate expected pupil progress.

> *'One of the ways in which we're accountable if Ofsted come in essentially they will talk to the children and they will look at the books. Yes, they will talk to the children, but the weighting is actually what they see in the books, do they see progression? It's a shame actually, because that in a sense threatens the principles of dialogic teaching'* (mentor, RegB-S5).

The mentor sought to quell these concerns from a school perspective, impressing that: 'you can't get the quality in books if you're not doing the discussion and the dialogic work that builds into that work'. It was felt over time this 'turned the tables a little bit' towards the teachers feeling 'validated' to employ Dialogic Teaching approaches, even if this did not necessarily result in anything being recorded.

Technical issues

Interviewees from two separate schools stated that the technology used to support the recording of lessons could have been more advanced and that it occasionally compromised the quality and the authenticity of video footage.

> *'I can see it's well thought through. There's better and easier technology to use now, like the IRIS technology. Sometimes the technology, and having to plug wires in and all that, can be a bit of a barrier. I think there's better stuff on the market for doing that that would potentially be more inspirational. You could use iPads and things like that; you could make it more accessible. I do think the technology is a bit archaic'* (headteacher, RegB-S4).

Finally, the mentor from School 8 in Region B stated that participants from their school had a strong preference for using audio recordings as opposed to videos:

> *'We've found, and we all have found this, that we much prefer the dictaphones to the videos, because there's no visual distraction. I much prefer closing my eyes and listening to the dictaphone conversations, both for the whole class and within groups and pairs and things to the video, because it's just talk, isn't it? It's just there, it's just sound. We've all said that separately, that it seems to be more advantageous to do that.'*

Is the intervention attractive to stakeholders?

This intervention was considered highly attractive, and a valuable experience, by stakeholders. All groups surveyed (headteachers, mentors, and teachers) reported that they would recommend the Dialogic Teaching approach to colleagues. The majority of teachers, in interviews and survey responses, reported that they planned to continue to use this approach in their teaching for the foreseeable future.

Fidelity

Was the intervention delivered as intended to all in the treatment group?

Of the 38 schools in the treatment group, six withdrew at a very early stage—five before the programme had started, and the sixth within three weeks of the beginning of the first term of teaching. All six schools gave reasons associated with staffing capacity relating to absences, or recent changes in staffing, that meant it was no longer possible for them to engage with the intervention.

Attendance at training sessions

Among the remaining 32 schools in the treatment group, there was a very high level of fidelity. Headteachers, mentors, and teachers in all schools participated in all training sessions. Where staff from a school were unable to attend a scheduled training session, a catch-up session was organised by the project development team. Data regarding attendance at training sessions was provided by the project team in the form of training session sign-in sheets, and by personal communication.

Mentor meeting fidelity

All teachers completed all required cycles of mentor meetings. This included a total of 11 meetings during the course of the intervention. The majority of schools (57%) completed the 11 cycles of mentor meetings on time, while the remainder completed them over a longer period of time (but still before testing for outcome measures).

If there were any issues with fidelity, what were the reasons?

No issues with fidelity were identified aside from those discussed in the previous section. These related either to time management, potential conflict with other agendas, the complexity of the intervention, or to technical issues related to use of the recording equipment. However, evidence from the survey, interviews, and case studies all indicates that teaching practices changed substantially as a result of schools' participation in the intervention, in line with the Dialogic Teaching approach. This is discussed in more detail below, under 'Outcomes'.

What elements of the intervention are perceived to be adaptable?

As discussed above, there were two main ways in which participants adapted the intervention. One was to adjust the timings of the mentor meetings which resulted in more than two weeks between some meetings. A second way was to focus on particular aspects of the Dialogic Teaching training materials in order to reduce the complexity of the intervention. This generally involved a focus on the nine 'talk-moves'—the majority of teachers reported that working towards increasing their repertoire of talk-moves during the course of the project was their primary aim.

In thinking about how the intervention could be managed beyond the duration of the project, headteachers from five schools described detailed plans for roll out, but with some variation from the intervention as implemented during the project. For example, in School 3, Region B there was commitment for the participating Year 5 teachers to continue using a Dialogic Teaching approach, although they expected the mentor meetings to cease. On the other hand, in School 4, Region B the headteacher planned to use the video recording of lessons more widely—with subsequent peer discussion and reflection—with less emphasis on Dialogic Teaching:

> *'I think the principles of getting teachers to maybe use video to reflect, to have the confidence to watch it, to share with a colleague, to be able to spot the bits and change their practice, you know, I think we will be looking to take those elements forward'* (headteacher, RegB-S4).

> *'Well one of the things we've asked the three staff who've been involved in the project to do is to do a presentation to the rest of the staff and to do some staff inset about their learning from the project and the things that have had the most impact. We have regular "sharing good practice" inset sessions where all staff are encouraged to share something that worked really well in their class. So they're going to be doing some of the strategies from dialogic teaching initially, so the teachers can try them out. Then next year in our inset programme we're going to include more of the key learning from the project, and then how we want to spread the practice throughout the school'* (headteacher, RegB-S8).

Outcomes

What are the perceived outcomes of the intervention?
Changes in classroom practice

Overall, participants reported relatively even usage of Dialogic Teaching across the three core subjects (maths, English, and science). In terms of other, non-core subjects, responses were less forthcoming, although history, geography, art and PSHE were mentioned as specific examples where the Dialogic Teaching had been used with varying degrees of perceived success. Respondents varied in their reporting of how well they had been able to incorporate a dialogic approach across different subject areas—there was an approximately even split of participants who reported the approach being easier in each of maths, English and science. A more nuanced exploration of the data suggests the more important predictor of perceived success in integration was related to the nature of a particular task, as opposed to a subject area. Lessons that involved more investigation or exploration were said to lend themselves well to this approach, as were other tasks that did not have a 'right or wrong answer'.

The case studies showed some variation in the degree to which classrooms had adopted a Dialogic Teaching approach. While this is a very small sample, it demonstrates the full implementation of this intervention was relatively complex, and that there was some variation in the extent to which this was achieved.

Participants made reference to enhanced teacher confidence and improved pedagogy on numerous occasions. These were sometimes linked with the implementation of specific features of the Dialogic Teaching approach itself (such as specific repertoires or talk-moves) and sometimes with more general

aspects of the project, such as reflecting on video observations with mentors. Teachers cited their more sophisticated use of questioning—particularly with regard to how follow-up questions were used more skilfully:

> *'What I've seen, particularly when I've done observation, has been the use of questioning and how teachers have used questions, but it's not just been about one question. So, they might have asked a question, the child has given an answer, so they might have a follow-up question and then an even further follow-up question which might go to the same child; it might go to a different child, rather than having a series of not particularly connected questions. So, it's more of a response to what the child says and then either encouraging the child to further explain, or maybe getting somebody else to add something to the explanation they've already given or something like that'* (headteacher, RegB-S8).

Specifically, teachers' introductions to lessons were often identified as having improved:

> *'I think the biggest effect it's had on my teaching is how I try and manage the introduction to a lesson, so it's more focused around how I can introduce things more quickly and getting the children working really, quicker than they would have done before'* (teacher, RegA-S2).

Interviewees were extremely positive about the impact on their practice of using videos as part of the programme—it became apparent many schools had not previously done so. Use of video, aligned to the dedicated space afforded at mentor meetings, helped teachers to improve their reflexivity and acted as a very powerful learning tool.

> *'At the start of it, I didn't realise I was so dreadful. I was giving them a couple of seconds to answer a question, which obviously wasn't enough. So, I think having that in the mentor meeting was good and that did make me reflect more day to day'* (teacher, RegA-S7).

Quantity of talk

In all cases, participants could give examples where quantity of talk had increased.

> *'The pupil talk has dramatically increased … They're bouncing off each other' (mentor, RegA-S7).*

However, as has been noted previously, dialogic talk was far more frequently used in the core subjects and many teachers contended that dialogic approaches lent themselves to particular types of lessons, for example those with a more exploratory focus and less narrowly focused on 'right' and 'wrong' answers. The impression given by some teachers was that Dialogic Teaching principles were employed for certain parts of a lesson as opposed to throughout, meaning that the quantity of talk was to some extent contained:

> *'It tends to be sort of small bits [Dialogic Teaching approach] in maybe each lesson. So, it might be the starter in a lesson, we'll try and plan in to do more of a discussion question, or to try to plan in opportunities to use the talk-moves and to have the children talking more. So, it tends to be a smaller bit of each lesson'* (teacher, RegB-S3).

Quality of talk

There was extensive feedback from participants about how involvement with the Dialogic Teaching programme linked to enhanced quality of pupil talk—often in quite profoundly positive ways.

> *'I think the quality of their talk has increased immensely and the way that they interact with each other over time has improved and they need less support to do it'* (teacher, RegA-S2).

Enhanced quality of talk was described in different ways but tended to focus around an increased range in the types of pupil talk that could be observed in the classroom without, or at least with reduced need for, teacher prompts or interventions.

> *'For me personally, as the programme has progressed through the different cycles, I'm seeing children more engaged and able to articulate things much better than they could before'* (mentor, RegB-S5).

> *'When you are talking to them, so in whole-class teaching situations, and when you are talking to them in smaller groups yes, you can see their answers are being extended naturally without that prompt'* (teacher, RegB-S3).

> *'I think the major thing, the important thing, is that the quality has increased. They're saying things with more purpose and more direction. As I say it sort of flows, there's more structure to a class conversation because they're all ... rather than the classic one where they're all sitting on the carpet with their hand in the air desperate to say their point that they thought of before the discussion started. I think they certainly last longer because the children talk for longer and there are more children involved for longer, rather than switching off'* (mentor, RegB-S8).

Several respondents reported that the Dialogic Teaching intervention afforded a language and a set of structures that created a 'safe environment' for pupils to speak openly about their views and to challenge the views of others. Over time, pupils arrived at a form of etiquette in the classroom that was conducive to improved quality of talk.

> *'I think it has [improved the quality of talk]. I think at the very beginning I don't think they were able to disagree respectfully, or have the courage to disagree with someone without being afraid to hurt their feelings. I think as we're going along they're realising that it is okay to have a different opinion, as long as we say it nicely. It is okay to say no, actually I don't agree with you, as long as we can justify it. It has definitely got better. They're using a lot more formal vocabulary, formal words: "I don't agree because I think..." and stuff like that. It's actually quite good'* (teacher, RegB-S8).

Some teachers reported different levels of impact for different categories of pupil. The following teacher thought that the dialogic approach was not equally successful across all students within their class, and reported that lower ability pupils found it more challenging:

> *'You do have the issue of engagement for certain children. Obviously, the ones who it's more challenging to keep them engaged in a lesson can find the discussion, the talk, a little bit more challenging, because you've got to have them interested in the first place. The lower ability, it can be quite hard for them to sometimes process what others are saying, especially if you're going at quite a pace in class. So yes, it has varied across the class. Some of my children you can see it's had more of an impact than others … I'm not going to lie and say all—but there's a small group in each class that are starting to use the question types when going round, and the discussion'* (teacher, RegB-S3).

Pupils with English as a second or other language (EAL) were also mentioned as a group that found Dialogic Teaching comparatively more difficult:

> *'It was challenging for the EAL children, because they haven't got the confidence to speak in English. I've had two children in the class throughout the year who've come with no English at all, so they struggle to have the confidence to speak. They're getting there now, but at first it was very difficult'* (teacher, RegB-S4).

However, the quotation below encapsulates the views of most participants who stated that, despite acknowledgement of variability in terms of how quickly, and to what extent, different types of pupil improved their quality of talk, pupils did all benefit as result of exposure to a Dialogic Teaching approach.

> *'I think all the children have improved, so I don't know whether the gap between the best and the least talkers has narrowed, but they've all got better, if you see my point. That's the issue with all the closing the gap things, that whatever you put in place helps everyone'* (mentor, RegB-S8).

The impact evaluation was not able to test differential effects for pupils according to difference in first language or prior attainment. However, these would be useful interactions to test for in a future evaluation.

Changes in classroom talk from the development team analysis

Results within this subsection come from the work of the development team (CPRT/University of York) and have not been scrutinised or validated by the independent evaluation team. The quoted text has been adapted from Alexander with Hardman (2017). "In all three core subjects, the ratio of closed to open questions was evenly balanced at the beginning of the project but by the end, intervention group teachers were making greater use of open questions than their control group peers (Appendix I, Tables 1–6). This implies a more dialogic approach in classrooms, with increased levels of pupil talk.

Intervention teachers were trained to deploy a variety of moves to probe, extend, and follow up pupil contributions on the principle that these would both improve engagement and yield cognitive gains. Differences between the two groups in respect of these were most marked in maths and science, where by phase 2 the intervention teachers were making significantly greater use of wait time, revoicing, rephrasing, seeking evidence of reasoning, challenging, requesting justification, and so on.

In English and maths, the ratio of brief to extended pupil contributions in September 2015 was the same in intervention and control classrooms. By February 2016, there were differences between the groups in respect of an increase in extended pupil contributions and a decrease in brief contributions.

In science, the intervention group started the programme with a higher ratio of extended to brief pupil contributions than the control group (given that this happened after the induction and training it may suggest that the programme's messages in this regard were more readily implemented in science than the other two subjects, or even that primary science teaching tends to be more instinctively dialogic). This difference was sustained through to the end of the intervention.

In English and maths, comparable ratios of recitation to discussion/dialogue in the intervention and control groups were transformed into differences by February 2016, with intervention teachers making much greater use of discussion/dialogue.

In this matter, science was again somewhat different in that at the beginning of the trial, the intervention group was already making greater use of discussion/dialogue than the control group. This continued through the project."

Attainment in tests

No respondents gave any indication that they felt the Dialogic Teaching programme would have an adverse effect on their pupils in terms of attainment or testing.

> *'I know it definitely hasn't done them any harm and I feel like the attainment will have increased due to the project. I don't think it will be dramatically increased—we haven't seen any dramatic improvements in attainment'* (mentor, RegA-S7).

However, participants were generally sceptical about there being improved attainment within the two-term scope of the project. This was not because of a lack of conviction in the programme or the principles of dialogic teaching but more because of their belief that the transition to a Dialogic Teaching approach was a long-term project, and that it would take more than two terms for such changes to influence attainment. The programme was perceived to be very intense and required a period of adjustment for both pupils and teachers. The quotation below was indicative of many participants' concerns and specifically advises that the programme should be a whole-school project starting from reception class to maximise the likelihood of improving attainment.

> *'Wouldn't have a clue is the honest answer to that [the likelihood of pupil attainment having increased], and I wouldn't expect it to in what is effectively a term and a half. I wouldn't expect it to have impacted. I think this is a long—I know the project is only for two terms—but I think this is long-term. I think it's actually something that needs to come through school from Reception and Year 1. I think a lot of it is skills that they need to adapt and build on as they go'* (mentor, RegB-S3).

Participants gave many positive examples of their belief in the theory underpinning the Dialogic Teaching approach. Most expected a positive impact on pupil learning and attainment over time. No more so than the following mentor:

> *'What we see in the classrooms and the interaction of the pupils and the way the teaching has changed, that can only have a positive impact on progress and therefore attainment. I've been teaching for 21 years and yes, if that doesn't have a positive impact, I don't know then what does—it's got to'* (mentor, Reg B-S5).

Finally, it is important to note that there were frequent examples of participants describing positive outcomes having been achieved, other than attainment in tests:

> *'I'm not convinced at the moment that pupil attainment will change. I think pupil confidence would come first'* (mentor, RegA-S1).

Confidence

There was strong evidence that overall pupils had greatly gained in confidence during the period of involvement with the programme, with no examples cited of pupils being negatively affected. Many participants reported a big improvement in pupils' willingness to 'have a go'—particularly among those previously more reluctant to engage. This enhanced engagement was attributed to the creation of a 'safe classroom environment' which enabled pupils to speak and to challenge each other:

> *'As I said, I do think the children, I must say in my class, are more able to discuss things. If I pose a question, at the start of the year we'd have quite a lot of silence. They wouldn't really be sure how to answer it, whereas now, even if they're not sure, they will give it a go, they will try'* (teacher, RegA-S6).

While improved confidence was a big positive, one teacher described a group of 'intelligent boys' that tended to 'take over' and spoke of the 'fine line' between embracing the natural enthusiasm of the most able pupils and not negatively impacting on the wider group. Finally, the example presented below powerfully illustrates how the mentor meetings and analysis of video footage helped one teacher identify negative practice around inviting only pupils they thought would provide the 'correct' answer. Through using a random name generator, a pupil who rarely volunteered to provide an answer was selected and proceeded to provide a 'really well-articulated answer' which helped boost their own confidence and the resolve of the teacher to invite a broader range of pupils to give answers in the future.

> *'He had had some peer talk, so he'd had time to think about the question, so to prepare, and they knew that a random name was going to come out, and he gave a really well-articulated answer. Really that was to be celebrated, because (a) the teacher said, and this is in the review, "I would never have normally chosen that child". It made her realise that often you choose children who you think are going to give the right answer that you want to hear, so it was really useful. Also, (b) that child had the opportunity to be involved, which for them was a great boost and obviously more opportunities then developed from that one experience. I would say those are the challenging groups of children. That doesn't necessarily mean we've not had our successes in them'* (mentor, RegB-S5).

Were there any negative outcomes or unintended effects?

The only negative outcomes observed during the process evaluation relate to the time management difficulties discussed above while teachers and mentors were engaged with the project.

Formative findings

Are there any ways that the intervention can be improved?

The clear majority of feedback was very positive. However, participants identified a small number of areas for improvement should the intervention be run again.

Streamlining the programme and making it more flexible

The most commonly cited improvement, as has been consistently reported throughout the process evaluation, was that the programme would benefit from being less intense, run over a longer period, and with greater flexibility woven into the design. Many responses indicated that making these alterations would be make for a better and more realistic structure for school implementation.

> *'I think the cycles are very, very tight, which has put the teachers—particularly with the videoing—under a bit of pressure …There's been no breathing space really to reflect: yes, you do this and then we're straight on to the next cycle. So, I suppose having a bit more breathing space really, a little bit more flexibility, because, for example, Christmas happens, doesn't it, and things go a little bit to pot, so one of the cycles went a little bit wrong there'* (mentor, RegA-S6).

> *'I think in an ideal world it would work well, but here it isn't an ideal world and things come up, so that's been our issue with it really … Well, in the current format for a school in our situation I don't think it's realistic in that time-scale with the expectation that things happen every fortnight'* (mentor, RegB-S4).

The following participant specifically argued for a 'less is more' approach to future delivery, claiming it would be far preferable to concentrate on a much smaller number of foci at any one time, thus making the model more manageable and likely to be delivered in a consistent and quality manner.

> *'Essentially that would be it, that less is more. In terms of the demands that were already made upon class teachers, really focusing on one quality thing at a time and making sure that is then implemented across in different aspects of the curriculum, rather than trying to think about lots of different things. Because the tendency then is if you've got too many things to think about you just shut down'* (mentor, RegB-S3).

More opportunities to share experiences with other schools

A further suggested enhancement to the programme mentioned by many participants was that there be greater networking opportunities for sharing experiences and good practice between participants from other schools. Although it was acknowledged that the school liaison officer had circulated participants' email addresses, it was generally held that more could be done to make networking opportunities a more formalised part of the programme. Specific suggestions included an event at the end of the programme, online forums, and school visits.

> *'I think we need more opportunity for the schools to liaise with each other. That's something that we haven't really had. It's very much been "this is our school project and this is what we've done". We haven't had an opportunity to discuss with other schools. Timing is one of the issues. When do you do that? Do you do it after school? Who's willing to give up more time? Schools find it very hard at the moment anyway to release people, so that's going to be an ongoing battle, isn't it?'* (headteacher, RegB-S5).

Duration of the intervention

As discussed above, under 'Outcomes', the majority of participants felt that the intervention had not lasted long enough for it to have maximum impact within their schools. Some respondents stated that they thought it would take two or three years for the new approach to 'bed-in' and for new classroom cultures to establish themselves. Some teachers stated that the full impact could only be appreciated if a Dialogic Teaching approach began in Reception class and continued through to Year 6. This issue is discussed in more detail below, under 'Limitations'.

Control group activity

What happened in the control group and how did this compare to what was intended?

We consider here control group activity, in particular whether there was any resentful demoralisation or compensation rivalry going on. We ask the question: what was 'business as usual'?

It was not possible, within the scope of this evaluation, to collect data on control group activity. This means that there is a risk that schools in the control group could have been using elements of a Dialogic Teaching approach, or could have been engaging in other development or intervention projects to raise attainment, possibly even to compensate for disappointment in being allocated to the control condition. This is a limitation of the evaluation. The development team did collect some data from a sample of schools in the control group (see above) which showed differences in classroom practice between the intervention and control groups (increased quantity and quality of pupil talk, and reduced quantity of teacher talk in intervention classrooms, compared with controls).

The control group had the opportunity to participate in training for Dialogic Teaching at the end of the first year. Those schools that took this opportunity were also able to use the video and audio recording equipment to support the development of a Dialogic Teaching approach. Training and equipment were offered to control schools on the condition that they would not use the approach with Year 6 pupils in 2016/2017. This condition was imposed so as to minimise the risk of contamination during the year

leading up to Key Stage 2 national tests, representing the secondary outcome measures for the impact evaluation.

Conclusion

Key conclusions	
1.	Children in Dialogic Teaching schools made two additional months' progress in English and science, and one additional month's progress in maths, compared to children in control schools, on average. The three padlock security rating means we are moderately confident that this difference was due to the intervention and not to other factors.
2.	Children eligible for free school meals (FSM) made two additional months' progress in English, science, and maths compared to FSM children in control schools. The smaller number of FSM pupils in the trial limits the security of this result.
3.	The intervention was highly regarded by headteachers, mentors, and teachers, who thought that the Dialogic Teaching approach had positive effects on pupil confidence and engagement.
4.	The majority of participating teachers felt that it would take longer than two terms to fully embed a Dialogic Teaching approach in their classrooms. It could therefore be valuable to test the impact of the intervention over a longer period.
5.	This intervention requires teachers to change classroom talk across the curriculum, supported by training, handbooks, video, and regular review meetings with peer mentors. Future research could aim to differentiate the effects of these different elements.

Limitations

Limitations to the methodological approach taken here mean that there may be positive effects of the intervention that could not be detected, or that effects observed may be underestimated.

Intervention duration

The intervention lasted only two terms. It is arguable that this was not enough time for the change in teaching approach to fully embed in teachers' practice. This intervention required a relatively large number of major changes in teaching practice, classroom management, and curriculum design, leading to major changes in pupil learning behaviour, leading to changes in pupil attainment. Not only this, but the project delivery team asked participating schools to implement the intervention across the curriculum, with a particular focus on English, maths, and science. Rossi, Freeman and Lipsey (1999) caution against evaluating complex interventions in their first year of implementation. Similarly, Ginsburg and Smith (2016) suggest that it can take two or three years for a complex intervention to be reliably implemented.

Data from the process evaluation supports the claim that the intervention did not last long enough to have the maximum effect on children's attainment. The majority of participants (including headteachers, mentors, and teachers) reported that the Dialogic Teaching approach was having a positive effect on pupils' learning, as well as on pupil engagement, confidence, and motivation, but that they did not expect to see increased attainment within the scope of the project. Many teachers told us that they thought that it would take more than a year for this approach to have an effect on attainment. Similarly, headteachers, when asked about their intentions to use this approach in the future, responded by saying that they thought it needed to start in Key Stage 1 and run throughout children's entire primary school experience for it to be fully effective.

Attrition

The recruitment target for the trial was not met (78 schools were recruited, compared with the target of 80 schools). In addition, seven schools from the intervention group did not provide post-test data (~18% of the sample). While this attrition has not affected balance between the intervention and control groups on observed factors to a great extent (including KS1 attainment, EAL, and FSM) there may have been

unobserved differences between withdrawn schools and those that remained in the trial that could have affected the findings.

The forthcoming follow-up to this report will include an analysis of Key Stage 2 national test scores for pupils in the trial. This will include pupils in all schools that were allocated to the intervention and control groups in this trial. This analysis will address the issue of balance. However, this analysis could underestimate the effect of the intervention as it will include schools that dropped out of the intervention, and even those schools who remained engaged during the first year of the intervention may not continue to use a Dialogic Teaching approach in the second year. An 'on-treatment' analysis will also be reported in the addendum that includes only the 31 schools in the control group that participated fully in the intervention.

Control group activity

No process evaluation data was collected from the control group so it is not possible to define what 'business as usual' meant for these schools. It is possible that disappointment at not receiving the intervention could have affected progress in these schools (either positive or negative). It is also possible that participation in the trial could have encouraged teachers to explore aspects of dialogic teaching themselves and incorporate this in their classroom teaching. Data collected by the development team ameliorate this concern somewhat as they show that classroom practice in intervention schools was different to practice in control schools at the end of the trial (more and higher-quality pupil talk, and reduced teacher talk in intervention schools).

Interpretation

There is evidence of a positive effect on children's attainment as a result of schools participating in this intervention. However, there are some indications to suggest that a longer intervention could lead to more substantial increases in attainment in national tests.

We can have some confidence in the positive effect observed in science attainment, and, to a slightly lesser extent, in English. For those pupils eligible for free school meals, there were consistently positive effects across outcomes and these were statistically significant or close to significant in some cases. These are promising findings, especially given the limitations described above, and suggest that there may be some value in carrying out an effectiveness trial over a longer period of time.

The process evaluation showed that the intervention was universally well received by schools. Participating teachers, mentors, and headteachers believed that the intervention was of benefit to children's learning. More specifically, teachers believed that the intervention had an effect on children's confidence, their engagement, and their ability to reason effectively. Teachers generally thought that more time would be required in order for the Dialogic Teaching approach to have an effect on attainment in national tests.

This was a complex intervention. For the majority of teachers involved, it required substantial changes in practice, across the whole curriculum. This is quite different to the majority of intervention studies, which tend to focus on changing practice, and thus raising attainment, in one specific subject area. Teachers varied considerably in terms of how well-suited they considered the Dialogic Teaching approach to be for different areas of the curriculum. Some thought it worked well in science and English, for example, while others found it difficult to use in their English teaching but found it well-suited to maths. Some of these differences appear to be due to teachers' preconceptions relating to teaching and learning in these subjects (the extent to which a subject tends to have 'right' and 'wrong' answers, for example), and to perceived requirements from senior leaders and school inspectors. This would be a useful focus of future research as there was not scope within the current project to fully explore this finding.

A further complication relates to the fact that, as well as changes to classroom practice, the intervention involved a series of fortnightly mentor meetings focused on reviewing segments of video-recorded lessons. Many participating teachers felt that these meetings were the most valuable aspect of the intervention. Given existing evidence for the value of this kind of support (in the form of video interaction guidance, for example, as in Kennedy, Landor and Todd, 2011), it is not surprising that participating schools found it to be an essential component of this intervention. The design of the present trial did not allow us to dissociate the effects of changes in classroom practice from the effects of introducing mentoring for participating teachers. In a future study, it might be possible to test this with a three-arm design, including a mentoring-only group alongside the intervention and control conditions.

References

Alexander, R. J. (2001) *Culture and Pedagogy: international comparisons in primary education*, Oxford: Blackwell.

Alexander, R. J. (2003) *Talk for Learning: the first year*, Northallerton: North Yorkshire County Council. http://www.robinalexander.org.uk/docs/NYorks_EVAL_REP_03.pdf

Alexander, R. J. (2005a) *Teaching Through Dialogue: the first year*, London: Barking and Dagenham Council. http://www.robinalexander.org.uk/bardagreport05.pdf

Alexander, R. J. (2005b) *Talk for Learning: the second year*, Northallerton: North Yorkshire County Council. http://www.robinalexander.org.uk/docs/TLP_Eval_Report_04.pdf

Alexander, R. J. (2015) 'The CPRT/IEE Dialogic Teaching Project, trial stage 2015-16: handbook for schools', University of York (unpublished).

Alexander, R. J. (2015a) *Towards Dialogic Teaching: rethinking classroom talk* (4th edn), York: Dialogos.

Alexander, R. J. (2015b) 'Dialogic teaching and the study of classroom talk: a developmental bibliography'. Accessed from: http://www.robinalexander.org.uk/wp-content/uploads/2012/10/Alexander-dialogic-teaching-bibliography1.pdf

Alexander, R.J. (2017) *Dialogic Teaching in Brief*, http://www.robinalexander.org.uk/wp-content/uploads/2012/10/Dialogc-teaching-in-brief-170622.pdf

Alexander, R. J. with Hardman J. (2017) 'In-House Evaluation of the CPRT/UoY Dialogic Teaching Project', interim report, University of York.

Dawes, L., Mercer, N. and Wegerif, R. (2004) *Thinking Together: a programme of activities for developing speaking and listening*, Birmingham: Imaginative Minds.

Department for Education (2014) 'National curriculum in England: English programmes of study'. Retrieved from: https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study/national-curriculum-in-england-english-programmes-of-study

Ginsburg, A. and Smith, M. (2016) 'Do randomised controlled trials meet the gold standard?', American Enterprise Institute. Available at: http://www.aei.org/wp-content/uploads/2016/03/Do-randomized-controlled-trials-meet-the-gold-standard.pdf

Gorard, S., Siddiqui, N. and Huat See, B. (2016) 'Philosophy for Children: Evaluation report and executive summary', London: Education Endowment Foundation.

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D. [...] and Lamb, S. E. (2014) 'Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide', *British Medical Journal*, 348, g1,687.

Kennedy, H., Landor, M. and Todd, L. (2011) *Video Interaction Guidance*, London: Jessica Kingsley.

Lefstein, A. and Snell, J. (2014) *Better than Best Practice: developing teaching and learning through dialogue*, London: Routledge.

Mercer, N. (2000) *Words and Minds: how we use language to think together*, London: Routledge.

Mercer, N., Dawes, L. and Staarman, J. K. (2009) 'Dialogic teaching in the primary science classroom', *Language and Education*, 23 (4), pp. 353–369.

Mercer, N., Dawes, L., Wegerif, R. and Sams, C. (2004) 'Reasoning as a scientist: ways of helping children to use language to learn science', *British Educational Research Journal*, 30 (3), pp. 359–377.

Mercer, N. and Littleton, K. (2007) *Dialogue and the Development of Children's Thinking: a sociocultural approach*, London: Routledge.

Mercer, N. and Sams, C. (2006) 'Teaching children how to use language to solve maths problems', *Language and Education*, 20 (6), pp. 507–528.

Nystrand, M., Gamoran, A., Kachur, R. and Prendergast, C. (1997) *Opening Dialogue: understanding the dynamics of language and learning in the English Classroom*, New York: Teachers College Press.

Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S. and Long, D. A. (2003) 'Questions in Time: Investigating the Structure and Dynamics of Unfolding Classroom Discourse', *Discourse Processes*, 35 (2), pp. 135–198.

Lyle, S. (2008) 'Dialogic teaching: Discussing theoretical contexts and reviewing evidence from classroom practice', *Language and education*, 22 (3), pp. 222–240.

Pocock, S. J. and Simon, R. (1975) 'Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial', *Biometrics*, pp. 103-115.

Raudenbush, S. W. *et al*. (2011) *Optimal Design Software for Multi-level and Longitudinal Research* (Version 3.01) [software]. Available from www.wtgrantfoundation.org.

Resnick, L., Asterhan, C. and Clarke, S. (eds) (2015) *Socializing Intelligence Through Academic Talk and Dialogue*, Washington DC: AERA.

Resnick, L., Michaels, S. and O'Connor, C. (2010) 'How well structured talk builds the mind', in R. Sternberg and D. Preiss (eds) *From Genes to Context: new discoveries about learning from educational research and their applications*, New York: Springer.

Reznitskaya, A. (2012) 'Dialogic teaching: Rethinking language use during literature discussions', *The Reading Teacher*, 65 (7), pp. 446–456.

Reznitskaya, A. and Gregory, M. (2013) 'Student thought and classroom language: Examining the mechanisms of change in dialogic teaching', *Educational Psychologist*, 48 (2), pp. 114–133.

Rossi, P., Freeman, H. and Lipsey, M. (1999) *Evaluation: A systematic approach*, Thousand Oaks, CA: Sage.

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, pp. 734–739.

Wragg, E. C. and Brown, G. (1993) *Explaining*, London: Routledge.

Wragg, E. C. and Brown, G. (2001) *Questioning*, London: Routledge.

Zimmerman, P. H., Bolhuis, J. E., Willemsen, A., Meyer, E. S. and Noldus, L. P. (2009) 'The Observer XT: A tool for the integration and synchronization of multimodal signals', *Behavior Research Methods*, 41 (3), pp. 731–735.

Appendix A: EEF cost rating

Cost ratings are based on the approximate cost per pupil per year of implementing the intervention over three years. More information about the EEF's approach to cost evaluation can be found here. Cost ratings are awarded as follows:

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

Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	**Power**	**Attrition***					
5 🔒	Well conducted experimental design with appropriate analysis	MDES < 0.2	0-10%					
4 🔒	Fair and clear quasi-experimental design for comparison (e.g. RDD) with appropriate analysis, or experimental design with minor concerns about validity	MDES < 0.3	11-20%			Adjustment for Balance **[0]**		
3 🔒	Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%	3 🔒	→		→	3 🔒
2 🔒	Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%			Adjustment for threats to internal validity **[0]**		
1 🔒	Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	51-50%					
0 🔒	No comparator	MDES > 0.6	>50%					

- **Initial padlock score:** lowest of the three ratings for design, power and attrition = 3 padlocks
- **Reason for adjustment for balance** (if made): no adjustment made
- **Reason for adjustment for threats to validity** (if made): no adjustment made
- **Final padlock score:** initial score adjusted for balance and internal validity = 3 padlocks

*Attrition is measured at the pupil level, even for cluster trials.

Appendix C: Memorandum of understanding with participating schools

Dear head teacher,

CPRT/IEE Dialogic Teaching Project – Trial Phase

We are writing to invite you to participate in the main trial phase (2015-16) of the project *Classroom talk, social disadvantage and educational attainment: raising standards, closing the gap*, otherwise known as the CPRT/IEE Dialogic Teaching Project.

The project, which is funded from 2014-17 by the Educational Endowment Foundation (EEF), entails collaboration between the Cambridge Primary Review Trust (CPRT) and the Institute for Effective Education (IEE) at the University of York, where the project is based. Its joint directors are Professor Robin Alexander (CPRT) and Professor Frank Hardman (IEE). Like all EEF projects, this one aims to develop and trial a practical strategy for raising standards of attainment among pupils, especially disadvantaged pupil. In this case, the strategy is a variant on dialogic teaching, an approach initiated by Robin Alexander, piloted in several local authorities (including Barking and Dagenham) and now used or adapted in many schools in the UK and other countries. This prioritises high quality classroom talk as a tool for enhancing pupils' motivation, engagement, participation, thinking and understanding, and hence raises the standard of their learning. There is now good evidence that properly pursued, the approach achieves these goals, but as yet there has been no UK study which uses the rigour of a randomised controlled trial to settle the matter definitively. This is what our project aims to do.

The project has two phases. The development phase took place in 2014-15, during which the strategy was piloted and refined with Year 5 teachers in ten primary schools in Barking and Dagenham, London. Barking and Dagenham is a local authority with considerable collective experience of the approach in question and from whose experience and insight the project greatly benefitted.

The randomised controlled trial (2015-16) that we are asking you to participate in will take place in 80 primary schools in Birmingham, Bradford and Leeds with no prior involvement in Dialogic Teaching. As in the pilot, we work intensively with your Year 5 teachers on the Dialogic Teaching approaches during the autumn and spring terms of 2015-16. In addition to the CPRT/IEE team based at the University of York working with you and your teachers to study changes in the quality of classroom talk arising from the Dialogic Teaching intervention, an independent evaluation of its impact on learning outcomes will be conducted by the Centre for Education and Inclusion Research, Sheffield Hallam University. Assuming success, the approach will then be scaled up for national dissemination.

We very much hope that you will agree to participate. This is what the trial phase of the project will entail:

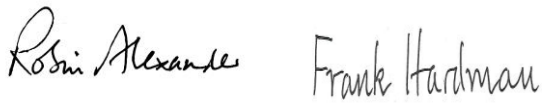
- School being randomly allocated to an **'intervention'** or **'control'** group.

- Schools in the **control group** will be asked to carry on teaching their classes as normal during the autumn and spring terms and will have the option of receiving the training in the Dialogic Teaching approaches in the summer term.
- All Year 5 children participating in the intervention and control schools will be tested at the end of the study by Sheffield Hallam University to study the impact of the intervention on learning outcomes using literacy, numeracy and science tests. Data on the pupils' Key Stage 1 and Key Stage 2 results in English and mathematics will also be obtained by the independent evaluation team from the National Pupil Data base.
- Each **intervention school** will appoint a further member of staff with relevant experience and expertise to act as the project teachers' mentor. NB: this will a collaborative relationship, from which all parties learn, not one of trainer/trainee.
- In July 2015 participating teachers, mentors and heads from all **intervention schools** will be invited to attend an **induction and training day** at which the programme will be explained in detail and the approach to development will be simulated in a practical workshop. Mentors will attend an additional training following on from the induction and training day in September 2015.
- The involvement of head teachers is vital, both to support the teachers involved and to make the Dialogic Teaching intervention a genuinely whole-school one in which ideas, suggestions and reactions are freely shared in order to maximise its impact.
- Thereafter, through a specified programme of video, review and refocusing sessions spread over the first term, teachers and mentors in the **intervention schools** will jointly work on targeted aspects of classroom talk.
- A random selection of teachers from the control and intervention schools will be video recorded at the start and end of the intervention to study changes occur in the patterning of classroom interaction and classroom talk. There will also be monitoring/support visits from members of the York team.
- At the end of the first term there will be plenary meeting of all **intervention school** participants to review progress.
- The process will be repeated, more or less, during the spring term.
- To guide and support the Dialogic Teaching intervention, every participant in the **intervention schools** will be given a purpose-designed professional guidance and support pack containing study and reference materials and a development and mentoring manual. The training materials will be made available to teachers in the **control schools** in the summer term.
- **Intervention schools** will be loaned all necessary equipment for the video recording during the autumn and spring terms which is an essential part of the process. **Control schools** will be able to access the video recording equipment in the summer term.
- Cover will be paid to the **intervention schools** for the induction/training day, mentor training day, the plenary days in terms 1 and 2, and an agreed number of teacher-mentor review/refocusing meetings. Two days' cover will also be provided to the **control schools** during the summer term to cover the induction and training day and mentor training.

This agreement letter provides a bare outline of the study. Further details will be provided during the training workshops in Birmingham and Leeds to be held the first half of July. In the next couple of weeks, we confirm the **intervention schools** and invite them to one of the induction days for which we will provide supply cover.

We hope that you are happy to sign and return the slip below to confirm your agreement to take part in the main trial on the basis outlined above, and that you recognise that the benefits of participation are mutual. With your help the project will develop an effective working strategy for maximising the effectiveness of talk for learning and teaching from which many other teachers and pupils will benefit; while your staff will engage in a professional development process from which they not only gain directly but which can be replicated and built on in future years and in other schools.

With best wishes,



Robin Alexander and Frank Hardman

Project directors

CPRT/IEE Dialogic Teaching Project

School Agreement Form – Main Trial

Name of school:

Name of head teacher:

We would like to take part in the CPRT/IEE Dialogic Teaching Project main trial during the 2015-16 school year on the basis outlined above.

As head teacher I confirm that, in discussion with my staff, we have agreed to participate in the main trial of the study. We are aware that we will be randomly assigned to either the intervention or control group. If selected to be in the **intervention group** we agree to our Year 5 teachers receiving the intervention as outlined above during the autumn and spring terms of 2015/16.

We agree to seek the permission of Year 5 parents/guardians to participate in the testing and video recording of classes as an intervention or control school. We will use the opt-out form provided by the CPRT/IEE delivery team giving permission for their children to be involved in the study, to be tested at the end of the year, and for the evaluation team to access their Key Stage 2 SAT results in English and Mathematics from the National Pupil Data base in 2017.

We also agree to provide the delivery and evaluation teams with the following information:

- Year 5 class lists
- Names, dates of births and Unique Pupil Numbers (UPNs) for these 2015/16 Y5 pupils;
- Names of pupils with EAL
- Names of pupils eligible for FSM
- Details of teachers who will be involved in the dialogic teaching intervention (names, job title, qualification year)
- Details of the mentor for the school (name, job title, qualification year)
- Access to video and audio recording conducted for professional training purposes
- Teacher surveys carried out in the spring term
- A list of children whose parents have opted out of the study

Signed on behalf of school:

Head teacher: Chair of Governors:

Date:

Please detach, sign and return this slip as soon as possible to:

Dr Taha Rajab

Institute for Effective Education, Berrick Saul Building, University of York, York YO10 5DD

Appendix D: Letter to parents

Dialogic Teaching: Improving talk for teaching and learning

Dear Parent/Guardian,

We would like to request your permission for your child to take part in an educational research project known as 'Dialogic Teaching' funded by the Educational Endowment Foundation (EEF) set up by the Department for Education. The teacher development project involves a partnership between your child's school, the Cambridge Primary Review Trust (CPRT) and the Institute for Effective Education (IEE) at the University of York. In addition to the CPRT/IEE team, an independent evaluation of the impact of dialogic teaching on learning outcomes will be conducted by the Centre for Education and Inclusion Research, Sheffield Hallam University.

The following information explains why the project is being conducted and what it would involve for your child.

What is the purpose of this study?

This study is being carried out to investigate the link between the quality of teacher-pupil talk and the raising of learning achievement in English, mathematics and science.

Why is my child's school participating?

Your head teacher has invited us to the school to work with teachers in Year 5 on a training programme to help them improve the quality of classroom talk to help raise learning outcomes in English, mathematics and science.

What will happen in the study?

Teachers will study the way they talk with children when they ask questions and follow up answers from the children.

Your child's school will be randomly allocated to one of two groups: an '**intervention group'** where teachers will receive the training in the autumn and spring, and a **'control group'** where teachers will receive the training in the summer term.

We will also access your child's Key Stage 1 and Key Stage 2 English and mathematics SATs results from the National Pupil Data Base in 2017 to see if the training programme has had a longer term impact on your child's learning. It will be shared with the EEF and UK Data Archive for research purposes only. At no point will your child be identified in the research and analysis will always be at the school level.

By delaying the training of teachers in the **control group** until the summer term, we will be able to see what differences, if any, the training in classroom talk is making to the quality of teaching in the **intervention schools.** It will also enable us to make improvements to the programme before it is delivered to the **control schools** in the summer term.

When receiving the training, your child's Year 5 teacher will work with other teachers in the school to video record and review the way they talk with children in whole class, group-

based and one-to-one situations during English, mathematics and science lessons. Some recorded lessons will also be analysed by a team of researchers based at the University of York.

What will your child's participation be?

Children will participate as normal in English, mathematics and science lessons regardless of whether they are in the **intervention** or **control** group of schools. As discussed above, children will be assessed by the Centre for Education and Inclusion Research at Sheffield Hallam University using English, mathematics and science tests at end of the year and the results compared with Key Stage 1 scores. Some of the lessons in the intervention schools and control schools will also be recorded for research purposes in order to analyse the quality of the talk that goes on between the teacher and children.

What should you tell your child about the study?

It would be helpful if you could tell your child that the research study is trying to find out how their teacher talks with children during English, mathematics and science lessons so as to make them more interesting and to help in their learning.

What are the possible benefits of taking part?

By participating in this study your child will help us study how the quality of teacher-pupil talk can be improved to help raise attainment in English, mathematics and science.

What happens when the research stops?

Researchers will analyse the data from the video recordings and assessments of English, mathematics and science and link them to the wider assessments that occur in schools. Scores for individual pupils and classes will be shared with the school but otherwise will remain confidential.

Will my child's information be kept confidential?

Yes. Data will be securely stored in line with Data Protection requirements. All assessments will be anonymized. Identifying data will be stored at the Institute for Effective Education at the University of York for a maximum 6 months for the purposes of data linkage, but identifying data will not be stored on the individual assessment forms. Any external reporting of the results will not identify any pupil or the school by name.

Parents may request that their children's data is withdrawn from the project at any point prior to identifying data being destroyed (i.e. within six months of the project start). They can do so by contacting Dr Jan Hardman at the University of York (details below). The data will be retained anonymously after the end of the project and may be used for future analysis and to link to other studies of a similar nature.

Does my child have to take part?

No. We are requesting your permission for your child to participate in the study. If you are **not happy** for your child to participate in the testing and video recording of lessons for the

research study, please complete and sign the attached **opt-out form**. A pupil's right not to participate in the study will be respected.

Dialogic teaching: Improving talk for teaching and learning

Parent/Guardian opt-out form

If you are not happy for your child to participate in the testing and video recording of lessons for the research study, please complete this form and return it to your child's teacher within the next week. (Please print clearly)

I do not give my permission for my child to take part in the research project.

Pupil's name: .....

Teacher's Name: .....

Parent's/Guardian's name: .....

Parent's/Guardian's signature: .....

Date……

In the case of any queries or complaint about the conduct of this study, please contact:

Dr Jan Hardman	or	Dr Emma Marsden
Department of Education		Chair of the Education Ethics Committee,
University of York, YO10 5DD		Department of Education
Tel: 01904 323499		University of York, YO10 5DD
Email: **jan.hardman@york.ac.uk**		Tel: 01904 323335
		Email: **emma.marsden@york.ac.uk**

Appendix E: Baseline Characteristics at Randomisation

Appendix E, Table 1: School Characteristics at randomisation (Intervention 38 schools, Control 38 Schools)

Variable	Intervention group		Control group	
School-level (categorical)	n/N (missing) / Percentage		n/N (missing) / Percentage	
School type				
LEA Maintained	31/38	82%	30/38	79%
Academies	7/38	18%	8/38	21%
Ofsted rating	37/38 (1)	97%	37/38 (1)	97%
Outstanding	5/37	14%	5/37	14%
Good	24/37	65%	25/37	67%
Requires improvement	7/37	19%	7/37	19%
Inadequate	1/37	3%	0/37	0
Eligible for FSM	38/38	35%	37/38 (1)	35%
EAL	38/38	52%	37/38 (1)	47%
Pupils with SEN Support	38/38	16%	38/38	16%
Gender				
Female	38/38	49%	38/38	49%
Male	38/38	51%	38/38	51%

Appendix E, Table 2: Pupil Characteristics at randomisation (Intervention 38 schools, Control 38 Schools)

Pupil-level (categorical)	n/N (missing)	Percentage	n/N (missing)	Percentage
Number of Y5 pupils	2492 /4958 (0)	50.3	2466 /4958 (0)	49.7
Eligible FSM	1256 / 2492 (0)	50.4	1221 / 2466 (0)	49.5
Gender				
Female	1209 / 2492 (0)	48.5	1227 / 2466 (0)	49.8
Male	1283 / 2492 (0)	51.5	1239 / 2466 (0)	50.2
Pupil-level (continuous)	**n (missing)**	**Mean**	**n (missing)**	**Mean**
NPD KS1 score	2,343 (149)	14.652	2340 (126)	14.815
NPD KS1 score (Ever FSM)	1205 (51)	14.201	1171 (50)	14.331
NPD KS1 score (Not FSM)	1138 (98)	15.130	1169 (76)	15.299

Appendix F: Model Covariates

Appendix F, Table 1: All Pupils Model covariates

GL English	Intervention			Intervention & KS1 Attainment			Pupil Background			School Minimisation		
	b	**s.e**		**b**	**s.e**		**b**	**s.e**		**b**	**s.e**	
Intervention	0.55	0.54		0.84	0.48		0.84	0.47		0.91	0.47	
KS1 (centred)	-	-		1.21	0.04		1.20	0.04		1.19	0.04	
EverFSM	-	-		-	-		-0.41	0.27		-0.36	0.27	
Sch. EverFSM	-	-		-	-		-	-		-0.05	0.02	
Sch. EAL	-	-		-	-		-	-		0.00	0.01	
Sch. KS2	-	-		-	-		-	-		-0.01	0.02	
Constant	**13.15**	**0.37**		**12.96**	**0.32**		**13.17**	**0.35**		**15.26**	**1.83**	
		s.e	**vpc**		**s.e**	**vpc**		**s.e**	**vpc**		**s.e.**	**vpc**
L 3 var. estimate	2.80	0.90	0.07	2.36	0.66	0.12	2.28	0.65	0.10	2.10	0.61	0.10
L 2 var. estimate	0.60	0.75	0.02	0.73	0.49	0.03	0.70	0.48	0.03	0.63	0.47	0.03
L 1 var. estimate	33.9	1.43	0.91	19.2	0.83	0.86	19.2	0.83	0.87	19.2	0.84	0.88
Total	37.3			22.3			22.2			59.5		
Total change variance / power	**0.07**	**0.2%**		**15.05**	**40%**		**15.08**	**40%**		**15.33**	**41%**	

GL Maths	Intervention			Intervention & KS1 Attainment			Pupil Background			School Minimisation		
	b	**s.e**		**b**	**s.e**		**b**	**s.e**		**b**	**s.e**	
Intervention	0.32	0.91		0.85	0.64		0.83	0.63		0.95	0.64	
KS1 (centred)	-	-		2.11	0.07		2.07	0.07		2.06	0.07	
EverFSM	-	-		-	-		-1.77	0.04		-1.61	0.45	
Sch. EverFSM	-	-		-	-		-	-		-0.05	0.03	
Sch. EAL	-	-		-	-		-	-		-0.01	0.01	
Sch. KS2	-	-		-	-		-	-		0.01	0.03	
Constant	20.94	0.62		20.50	0.44		21.41	0.49		22.60	2.51	
		s.e	**vpc**		**s.e**	**vpc**		**s.e**	**vpc**		**s.e.**	**vpc**
L 3 var. estimate	6.53	2.82	0.06	2.61	1.43	0.04	2.44	1.43	0.04	**2.16**	**1.41**	**0.04**
L 2 var. estimate	4.40	2.56	0.04	2.70	1.49	0.04	2.75	1.50	0.05	**2.94**	**1.53**	**0.05**
L 1 var. estimate	101.1	4.17	0.90	55.58	2.36	0.88	54.9	2.33	0.91	**54.4**	**2.23**	**0.91**
Total	112.0			60.9			60.1			**59.5**		
Total change variance / power	**0.02**	**0.0%**		**51.17**	**46%**		**51.96**	**46%**		**52.52**	**47%**	

GL Science	Intervention			Intervention & KS1 Attainment			Pupil Background			School Minimisation		
	b	**s.e**		**b**	**s.e**		**b**	**s.e**		**b**	**s.e**	
Intervention	0.42	0.77		0.67	0.55		0.70	0.54		0.96	0.46	
KS1 (centred)	-	-		1.50	0.05		1.48	0.05		1.44	0.05	
EverFSM	-	-		-	-		-1.21	0.35		-1.00	0.36	
Sch. EverFSM	-	-		-	-		-	-		-0.08	0.02	
Sch. EAL	-	-		-	-		-	-		-0.03	0.01	
Sch. KS2	-	-		-	-		-	-		0.03	0.02	
Constant	**26.26**	**0.52**		**26.29**	**0.38**		**26.89**	**0.41**		**28.39**	**1.84**	
		s.e	**vpc**		**s.e**	**vpc**		**s.e**	**vpc**		**s.e.**	**vpc**
L 3 var. estimate	6.38	1.88	0.09	2.85	0.95	0.07	2.62	0.91	0.01	**1.16**	**0.65**	**0.02**
L 2 var. estimate	0.60	1.23	0.01	0.55	0.77	0.01	0.52	0.76	0.01	**0.61**	**0.75**	**0.02**
L 1 var. estimate	60.99	2.53	0.90	35.63	1.52	0.91	35.4	1.51	0.92	**35.4**	**1.52**	**0.95**
Total	67.9			39.0			38.5			**37.2**		
Total change variance / power	**0.04**	**0.1%**		**28.96**	**43%**		**25.65**	**38%**		**30.75**	**45%**	

Appendix F, Table 2: Ever FSM Pupils only Model covariates

GL English	Intervention			Intervention & KS1 Attainment			School Minimisation		
	b	**s.e**		**b**	**s.e**		**b**	**s.e**	
Intervention	0.26	0.64		0. 76	0.57		0.71	0.57	
KS1 (centred)	-	-		1.10	0.06		1.09	0.06	
Sch. EverFSM	-	-		-	-		-0.05	0.03	
Sch. EAL	-	-		-	-		0.00	0.01	
Sch. KS2	-	-		-	-		0.00	0.02	
Constant	**12.62**	**0.32**		**12.74**	**0.39**		**14.07**	**2.18**	
		s.e	**vpe**		**s.e**	**vpe**		**s.e.**	**vpe**
L 3 var. estimate	2.70	1.24	0.07	2.81	0.95	0.12	2.51	0.862	0.12
L 2 var. estimate	0.39	1.25	0.01	0.03	0.73	0.01	0.00	0.00	0.00
L 1 var. estimate	33.89	2.13	0.92	21.47	1.38	0.88	21.29	1.294	0.89
Total	37.0			24.3			23.8		
Total change variance / power	**0.01**	**0.0 %**		**15.61**	**42%**		**13.2**	**36%**	

Appendix G: Case study findings

The case study findings are presented in relation to the six field questions outlined in the Method. Case study data are set out in Table 7 below.

Table 7: Sources of data for process evaluation in case study schools

School	Teacher Interview	Mentor Interview	Observations	Documentary Evidence
A	✓	✓	**Literacy, Maths**	
B	✓	✓	**Maths, English**	
C	✓		**English, Maths**	✓

The data generated through observations and video-recorded lessons reveal a range of responses to the dialogic teaching approach in classrooms. The discussion below discusses evidence of 'Indicators of Dialogic Teaching' (adapted from Alexander, 2015a) from the case study lessons, documentary evidence, and comments made by the teachers and/or mentors during interviews. Alexander (2015a) was provided to all teachers in intervention schools, and includes a set of 65 indicators. Case study researchers adapted this set by removing those items that were not relevant for their observations. For example, by removing items that relate to phenomena that could not be observed within a single lesson. Where numbers are given in parentheses, these refer to indicators in Appendix H (so for example, #33 stands for 'allows for range of responses [open / speculative &c.]'). Note that the set of indicators was used as a framework to guide observations, not as a measure of the extent to which dialogic teaching had been implemented in classrooms. The sample of 3 observed lessons was not enough to make a judgement about this. Rather, the purpose of the case studies, including the observations, was to explore some of the issues and the experiences encountered by teachers in the implementation of dialogic teaching in their schools.

1. ***What are the effects of the strategy for maximising the quality and educational impact of classroom talk using a dialogic teaching approach?***

The observed lesson in School A allowed students to build on previous knowledge (#31), and encouraged students to provide a range of answers (#33 & #34) in full-class exploratory talk. There was an environment of cooperation and all students were involved (#17, #18). In small-group discussions, students participated and worked on-task together (#22, #23), and often teacher monitoring made a difference to the talk (#27, #29). However, the lesson was still very teacher-centred, reducing chances for dialogue, and for most of the lesson the discussion was as a whole class. There were few opportunities for extended discussion (#13, #40) or for open discussion (#45, #46), with the teacher 'explaining' most of the students' contributions, and there were few opportunities for genuine pupil-pupil interaction (#22-#26), with the teacher controlling talk (cf. #22-26, #35 & 38); in a 55-minute lesson, more than 40 minutes were teacher-managed talk. The conditions for dialogic talk were not clearly in evidence (#1-4), and this lesson was not designed to assess students' oral skills (#11).

During interview, the teacher from School A recognised teacher-talk as an issue in her implementation of dialogic teaching and had discussed it previously with the mentor in development sessions. Despite this, she claimed to have noticed significant improvement in confidence and the level of participation of all students. She also noted that improved participation by all students was partly due to providing thinking time before demanding answers from students.

The teacher in School B started the sample lesson in a similar way to the teacher in School A, by asking students to build on previous work (#31), but quickly the repertoire of questioning expanded so that many responses were acceptable (#12, #14, #30, #32, #33, #35, #39, #45, #46). The responses from students noticeably built on other students' ideas (#22-24), often referring to other students by name (#17, #25). Many of the contributions were explanatory (#13, #15, #16), developing their own and others' ideas, and it was clear that listening to, responding to, and building on other students' responses had become the standard way to interact in discussions in this class (#3, #4, #9). Within the first few minutes of the class, the teacher encouraged as many students as possible (#16) to engage in expository talk, expressive talk and evaluative talk, providing some feedback where difficulties arose (#29, #42, #46). The teacher then had the students rearrange their seating, so that they were responsible for the talk in the class (#1, #2, #9, #10) creating better conditions for the development of dialogue by the students. In this phase, the teacher's main role was to nominate speakers who responded to each other's contributions to arrive at a shared goal (#7, #12, #15, #17, #19, #22, #23, #25, #39, #40).

The group worked together in a discussion that was instigated by the teacher but maintained by the pupils. From the very first exchange in this lesson, the pupils were trying to build upon each other's contributions (#22-24):

> *T: Right, can anyone remember, what have we been, what is the whole point of our objective this week? What have we been trying to do with our stories this week?... What have we been trying to do with our stories this week? P1?*
>
> *P1: We've been trying to show how a character's feeling, not tell.*
>
> *T: not tell, P1?*
>
> *P2: Like P1 said, erm, we've been writing, we've written, erm, we've stuck some little photos in and we've explained what they're doing but not using said.*
>
> *T: Not using the words. Describing. Excellent. P3?*
>
> *P3: So, we've been using vivid, vivid language, and er, by sticking in expressing how they feel but actually not telling the words like sad, happy*

Here pupil 2 uses the first pupil's name to add more information, while pupil 3 tries to summarise the previous contributions (#25). For the next 10 minutes, the teacher-led whole-group discussion provides many more examples of collaborative and creative talk, with most of the pupils adding non-predictable ideas to the discussion (#30-34). For instance, in one discussion of a story being constructed, one pupil commented "*I agree with P4, because P5 put a little too much adjectives in*."

This large group discussion took place without any teacher intervention or comment. In this lesson, students regularly disagreed with each other, but there was no indication of tension; students respected all other students' views (#26). This was one significant change that was noted by the mentor in her experience of the project classrooms:

A child can make a mistake and nobody's going to say 'that's wrong'. They're going to say, 'Well I understand what you're saying, but I think…'. School B Mentor

The mentor in this school recognised, however, that a key goal was to ensure that all students could take advantage of the learning opportunities offered by their dialogic teaching. The mentor was concerned that some students were benefitting more than others, and was therefore working on

development goals and techniques, such as nominating students more, to ensure that dialogic teaching worked to all students' advantage.

For the teacher in School C, dialogic teaching was understood somewhat differently than in schools A and B, as both the mathematics and English lessons followed a pattern where the teacher set up a task and then monitored, with students feeding back at the end of the task (#30, #37). The plan of this lesson provided potential spaces for a dialogic approach (#4, #6-10). Students spent most of their time in class engaged in transactional and exploratory talk during tasks and cooperating with each other (#6, #7, #17, #22, #23, #25, #32, #35, #38), and the teacher's monitoring of each small group was characterised by questions that scaffolded the process of learning (#27, #28). During feedback, students were constantly encouraged to offer reasons for their answers and to listen and respond to other students' answers with extended answers (#22, #24, #35, #45, #46). The teacher accepted all answers, made sure students listened to each other in order to engage in discussion, and typically only rephrased some of the contributions (#22, #43, #45).

The interview with the teacher revealed that the lesson plan, being largely based on tasks, was not influenced by dialogic teaching but was the approach that s/he would have used in previous years. As a result of their involvement in the project, this school introduced innovations such as assigning speaking roles (manager, time-keeper, summariser etc.) and talk-moves (asking questions to explain, clarify, extend, justify etc.) to students for each task. By the time of the case study visit students were familiar with these roles and they had an influence on most of the tasks in class. Both the mentor and the teacher commented that the speaking roles and talk-moves had eased and complemented the implementation of the dialogic approach. The teacher and the mentor mentioned independently in their comments that the dialogic teaching approach had proved most beneficial in science classes and least beneficial in mathematics, and suggested that where the lesson provided students with the chance to investigate and explore, the dialogic teaching approach had a lot to offer but in cases where the goal was a predetermined answer dialogic teaching was not as useful.

The teacher also felt, as in school B, that although the dialogic teaching project had benefitted all students, those who were already strong in class gained a greater advantage than those that were struggling to keep up with the curriculum.

In conclusion, Appendix H reveals that during the short visits made to the case study schools, evidence showed substantial implementation of dialogic teaching. Implementation across the three case study schools was noticeably uneven, but evidence from the observations and reporting by teachers and mentors all pointed towards an increase in the quality of teaching and in student talk.

2. How has an environment which fosters pupil attention and interest been promoted in the case study schools?

In school A, the teacher was positive about the approach and believed it had been beneficial in terms of pupil confidence and the opportunities to join in discussion. She felt the sentence openers had been useful in giving structure for speaking. She acknowledged that on viewing videoed lessons she noticed that "*I asked closed questions a lot so now I'm trying to make questions open and lead on from what kids suggest*". The observed lesson supports this view where there was much teacher talk and teacher questioning, about half of which was closed and in the hour lesson there were two opportunities for pupil to pupil talk - in the first instance for one minute and towards the end of the lesson for 5-10 minutes. The pupils were attentive to each other and the teacher throughout the lesson. The teacher was aware that while she may not always use a dialogic approach, she is "*very aware of teacher talk v pupil talk, I'm always conscious of it*". This sense of self-awareness, regarding questioning, has been arrived at through the sessions with the mentor, which the teacher found highly beneficial. It was through these sessions that the 'rules for talking' (which have been mounted on a poster on a board at the back of the room) were developed. The physical environment of the classroom was organised so that children were sitting on tables in groups of between 3 and 6. The teacher offered the chance for learners to change

where they were sitting but one child who was asked to move was told he could not sit with the partner he chose because he had misbehaved previously. During the observed lesson it was noticed that some of the children oriented their postures away from facing forward across the table and towards their talking partner when given the opportunity to talk in pairs. In these cases in particular the children were engaged and focussed on their talk partners.

In school B, the classrooms were arranged with the children sitting in groups in the Maths lesson and in a horseshoe shape with the other year 5 class having an English lesson: "*I have changed my classroom to a horseshoe and I much prefer it".* In the mathematics lesson the teacher could be seen asking a full range of question types, some inviting children to clarify their answers: "*Are you saying you agree with Martin?"* and some asking for evaluation "*would you rather split your delicious cake between 4 friends or 8?*" and some asking for explanations: "*Someone's just asked me how can I work out ¾ of 360? …how could I do that?"* At several points the teacher allowed the children to discuss amongst themselves across the classroom. What stood out was the respectful language the children were using to each other across the class; for example,

> *Girl: I agree with P1 but I have reason to not agree with him because you need more pieces to share it out equally.*
>
> *Boy 1: I agree with P2*
>
> *Boy 2: I disagree with P3 because the smaller the number, the bigger the fraction*
>
> *and in pair work:*
>
> *Girl 3: I agree with him… so both of them you up to 2 which is an even number…yeah but as well…*
>
> *Boy 3: Basically if you halve them all you get ones.*
>
> *Girl 3: yeah but that's a theory. But you are saying you get an odd number but you don't…*
>
> *and in the use of reasoning: "in fractions the bigger the denominator the less it is, you need you need two 1/8ths to make a quarter."*

Later the teacher pointed out one girl as being less good at maths, but during the lesson she was observed as able to explain her thinking: "*I think 8 is bigger than 4 so if you split them up into 8 pieces you'd have more".* This provides evidence that the classroom feels a safe place to make mistakes. Both teachers felt that the children benefitted from the approach and the teacher of the English lesson commented the children seemed to have "*a lot more confidence in speaking in class".*

In school C, the teacher thought that the dialogic teaching approach had influenced pupil engagement. They were now more prepared to listen to each other, and thus enjoyed their lessons more. The teacher identified the increase in group work (rather than paired talk) as the main change to previous practices. She indicated that clear guidelines on how to behave in group discussions had been beneficial:

> *I think they listen to each other a lot better now…(before) they'd all talk at the same time or they want to get their ideas across…now we're giving them the questions and particularly giving*

them the roles, they're very conscious of what they should be doing…listening to everybody else.

The teacher also explained that at the beginning of the year she had explained to the pupils what the project was about:

Hence, they know all about what I'm doing they know all about what I talk to K (mentor) about…because of that I've given them a lot of information and it's worked nicely because they then understand the project…they've been really involved.

One of the changes to the children's talk is that now "*they always give a reason…that's taken some training by me saying things like 'add on'*". The teacher feels that the class are "*quite good at using the talk-moves that have been provided*". The time for reflection upon the videoed lessons was seen by this teacher as formative in her promotion of talk in the classroom

3. What does the data set tell us about talk which meets and goes beyond the requirements of the national curriculum and what examples are there of pupils using talk to reason, explain, justify, argue, speculate evaluate and in other ways think for themselves?

The process evaluation team were able to gain a partial understanding of some of the ways pupils could be seen to be meeting or going beyond the requirements of the National Curriculum.

In each of the lessons in the three case study schools, most of the pupils could be described as "listening and responding appropriately to adults and peers", and "using questioning to peers and their teachers to extend their understanding and knowledge" (these and later similar statements in this section refer to extracts from Department of Education, 2014). In School A for example, a pupil asked "*Miss, did they have evacuees in the First World War?"* In the Maths lesson about fractions at School B, a pupil asked "*Do you do the same what you do to the bottom?"* to the teacher, and another pupil posed a question to peers in a small group: *"No, I'm talking about if you've got 8 friends, would you choose a quarter for yourself or share with eight friends?"*

In terms of "using relevant strategies to build their vocabulary", this appeared to be largely scaffolded by teachers, such as in School A where the teacher asked "*We've got four **eras.** Can anyone remember what an **era** is?*" Some pupils were observed articulating and justifying answers, arguments and opinions either in a whole class activity or in small groups or pairs. An example of this comes from School A where towards the end of the Literacy lesson the children were working in pairs cutting out pictures of toys and putting them in age order. Two girls had some discussion about ordering the pictures which demonstrates how their private conversation displays the incomplete sentences and deictic references (This, These) common to spoken English as well as (unelaborated) reasoning and justification.

G 1: Looks like from Victorian times.

G 2: this one's from today.

G 1: No.

G 2: look it's got…phones are now...then the next one this (rubics cube).

G 1: these ones, so, today for them.

G 2: The next one must be this one because…marbles?

G 1: because it doesn't look like a DS to me.

G 2: No, it can't go next…doing it from today to 1990's. So, the next one has to be marbles.

G 1: Next one has to be DS.

G 2: Yup

In all three of the case study schools, pupils participated in collaborative conversations, staying on topic and in a few cases initiating comments. In the literacy lesson at School A there were clear examples of pupils speculating, hypothesizing and imagining such as "*Could it be Ancient Egypt?* " and "*It looks like it was made in the Stone Age".* A further example was noted of a pupil exploring ideas: "*What is it with the brick cell phones? My mum told me about them…people had cell phones like bricks."* It is unsurprising that there were not more examples of extended reasoning or imagining hypothesizing and speculating when much of the teacher questioning observed required pre-determined answers, such as that seen in the literacy lesson in School B: "*You're going to be self-assessing. What do you think that might mean?"* And "*Why am I going to ask you to go over the success criteria? What do I look for when I am marking your work?"* A further example of this kind of questioning and the pupil responses from the literacy lesson in School B elucidates this point:

T: Can you see what techniques Pupil A uses?

P1: Use some sound effects

T: Yeah. Different types.

P2: Show not tell

T: Yeah building up tension. You need to….?

P3: Suspense?

T Yes and how do you create suspense?

P3: Short snappy sentences.

The pupils in School C were evidently comfortable speaking with peers in small groups on a variety of tasks, as well as contributing to class discussions by speaking "audibly and fluently with an increasing command of Standard English", either spontaneously or in reporting phases of the lesson, and in response to teacher-led questions. Evidence from the case study observations showed that, in terms of National Curriculum objectives, they could: "maintain attention and participate actively in collaborative conversations, staying on topic and initiating and responding to comments"; "use spoken language to develop understanding through speculating, hypothesising, imagining and exploring ideas"; and "speak audibly and fluently with an increasing command of Standard English." There was no evidence that "public speaking, performance and debate" was required for these students to extend their "confidence, enjoyment and mastery of language," as suggested by the National Curriculum.

4. What evidence is there of the dialogic teaching approach being used across the curriculum in the case study schools?

There was evidence of a dialogic approach to teaching in all classes observed, and in the video lessons provided by the schools, some of which are described below. These lessons included a specific focus

on literacy, mathematics and history. Interview data revealed that teachers' evaluation of a dialogic approach varied depending on the subject in the primary curriculum.

In the literacy lesson in school A, the teacher set both whole group and small group speaking tasks that contributed towards the overall aim of the lesson and enabled pupils to propose their own genuine answers, to challenge the teacher's opinions, and contribute vocabulary, although the majority of comment and follow-up was carried out by the teacher and some of the questions, particularly for short turns, were to display the correct knowledge of a predicted answer. During pair and small-group interaction, talk was primarily used to share and to reach a conclusion.

The history literacy lesson that was observed in School A included discussion of children, their living conditions and toys in different time periods. During the interview the teacher stated that a dialogic teaching approach had been most beneficial to the students when exploring themes in the curriculum, although little evidence was available to evaluate its impact in these other subjects. The teacher of this lesson described how the dialogic teaching approach had been used most in literacy and in topic- (or theme-) based parts of the curriculum. During the interview, the mentor revealed that mentoring sessions with the participating teachers had revealed greater success with lessons where students were investigating, particularly in science and literacy, rather than lessons based on 'factual retrieval' which the teacher reported occurring more often in mathematics classes.

In school B two different lessons by two different teachers were observed. The first was a mathematics lesson which began with a warm-up starter question which pupils discussed in pairs and then as a whole class group. This was followed by a fractions problem on a worksheet related to the colours of fireworks and the ingredient proportions required. The pupils worked on tables of 4 or 6 but mainly in pairs while the teacher circulated and posed questions to individual pupils to help them work out the problem. The questions to individual pupils tended to be closed and requiring one correct answer such as "what do you do when you simplify a fraction" and Is that an odd number?" and "What do you get if you divide 15 by 3?". When the teacher asked questions of the whole class however there was much more opportunity for questions requiring more extended answers such as "Does anyone still agree with P1?", "Are you saying you agree with P2?", "How can you test this out?" and "How could I do that?" The teacher of this class explained in the post observation interview her view that dialogic teaching was helpful for mathematics because "In the new maths [curriculum] they have to be able to explain and it's good for that." Her view was that dialogic teaching was more suited to mathematics, science, ICT and history than to literacy:

> *There's more discussion in maths than in literacy. It's so focussed on getting so much writing from them. Only once a week is there time for discussion of like small starters like teaching parentheses the other day. It was nice for them to say what they think it is. And science lends itself to dialogic questioning… In ICT we've been making a game and we discuss 'do you like what your partner's done?' rather than agreeing and disagreeing… And discussions in History, like about Henry VIII... 'would you have done things differently?'.*

The second observed lesson was a literacy class. The lesson also began with a warm up activity of a gap-fill where pupils were asked to identify the correct word class to put in a gap in a sentence. This was followed by individual evaluation of the pupils' own writing followed by peer review. This involved paired discussion of what worked well and what needed improving. The activity was followed by whole class discussion of the pupils' creative writing. The teacher was concerned about a perceived focus on writing for a certain time "*we've got to do 40 minutes of writing.*" Thus, "I*t's more difficult to do dialogic teaching in Literacy."* This teacher also felt that the teaching approach worked better in mathematics and science. When asked if she would continue with the approach, she answered "*Yes- but it doesn't work in every lesson"*. Having said that, she did go on to say "*I've changed my classroom to a horseshoe and I much prefer it* " and "*It's part of my teaching now… all my lessons have some sort of dialogic in them".*

The mentor at School B however disagreed with both teachers:

> *Yes, there is pressure of writing but I have seen good listening discussion. Literacy does lend itself. Numeracy lends itself only where they are doing investigations, not where something is being taught. Science does lend itself, such as the drawing of the moon lesson where all the drawings are different. School B Mentor*

In School C, during observed classes in mathematics and literacy, there were clear indications of a dialogic teaching approach. In an interview with the teacher, and in documentation provided by the mentor, care had been taken to attempt dialogic teaching approaches in the core subjects of the curriculum. The teacher's perception was that the approach worked best in science lessons and that literacy lessons also benefitted from dialogic teaching, but mathematics less so:

> *I think science it works really well with. I think that can be a real big discussion and you're exploring, you're investigating, you're trying things. We did a whole separating materials unit which it was amazing for, because I did a series of lessons where I gave them a range of equipment and they had to choose it and say why they wanted it and then they'd go and try something. If it didn't work they'd have to evaluate why, come back and change equipment. I think with practical lessons it's really strong. School C Teacher*

For instance, in mathematics classes the teacher explained that even when attempting to scaffold the process of tackling number problems expressed in words, many of the students were no longer interested in engaging in talk once they had identified the key aspects of the question:

> *He can explain that and he can give his reasons, but that's it then. That's where the discussion ends because they can agree or disagree, but when they know they've got the right answer it comes to a stop. School C Teacher*

Across the curriculum, the school worked hard on 'talk-moves': helping students to provide answers that explained, justified and gave reasons. Although the teachers focused on maths, science and literacy for this project, comments from the teacher and documented progress in the implementation of the project provided some evidence for these approaches to speaking being used by pupils in all subjects, including the non-core topics in the curriculum.

5. What are the teachers' and mentors' perceptions of the ability of the programme to raise standards in literacy, oracy and higher order thinking?

One teacher in School B believes this approach in her classroom is fostering independent thinking: "It's good to hear them arguing and understanding they could be wrong. Not giving them the answer but letting them go, until they get it themselves." The teacher of the mathematics lesson in School B believed there were advantages in terms of heightened communicative competence:

> *I've noticed how children will say 'I could be wrong.' They're listening to others, more in the Maths than in the English. There's less debate in English. They are more open to alternative viewpoints now… The children are willing to disagree, see another side. School B Teacher*

The NQT involved felt that their involvement in the project had raised her awareness of the role of talk in the classroom and that it had challenged previously held beliefs:

> *I think it's been instilled in me that if they're talking they're not working. So I get worried if they talk too much. I'm worried someone's going to come in and say why aren't they silent? I didn't have a clue what dialogism is when I started at the school. School B Teacher*

The teacher of the English lesson pointed out that although "We've got 40 minutes of writing. It's more difficult to do dialogic teaching in literacy," with a quiet class which doesn't "have ideas or join in much" the dialogic approach "brings them out of themselves" and so offers more opportunities to improve students' oral skills. She explained that before using this approach:

> *The children didn't join in - it was me speaking at them. They didn't come up with their own ideas… It does make me think about how differently different children think. One child last week suggested a tiger at the bottom of the lake. Before I would have dismissed it as silly, but when I asked her why, she gave me a good story about a tiger drowning. School B Teacher*

Perhaps the shift in teacher attitude due to the intervention has had the effect of creating a safe space for dialogue. "*They're a quiet class, but they don't have ideas or join in much so dialogic teaching brings them out of themselves. Yesterday there were lots of ideas creating creatures for their stories.*"

The teacher from School C stated that she believed that students' speaking and reasoning abilities had improved since the project started. The class have responded well to group discussions, talk-moves, and using different talking roles in the class, and they regularly gave reasons for answers, with or without prompts: "*There's always a <u>because</u> now*."

The teacher in school C was very concerned about the perceived ability levels of the pupils. Her concern is that "*most higher ability will listen and want to participate any way, whereas middle ability sometimes less so*". She felt that through using the talk-moves and encouraging giving a reason the children were more engaged. However, she felt that the dialogic approach "*is more difficult for lower ability to answer questions because they're quite higher order questions and higher order thinking, they have struggled to access and understand it*." The teacher felt that lower ability pupils struggled in group work over "*choosing and knowing when to use the moves appropriately*".

Furthermore, this teacher was certain that whilst the approach had worked relatively well with this class, because they "*aim to please, they are well-behaved*" she felt that her class the previous year, characterised by discipline problems, would not have been as amenable and "*we would have had to work a lot harder on actually engaging them to use the questions at all*. ".

6. What are the mentors' views of the effectiveness of this model of professional development, its sustainability and its potential to be scaled up?

The project itself was viewed very positively ("*fab*") by the mentor in School A who warned that "*the timeframe is tight*". The difficulty of timetabling a meeting time for all 3 year 5 teachers to meet to discuss planning for the project was referred to and the difficulty of timetabling specific lessons for 'doing dialogic teaching' was raised. The mentor in School A felt strongly that the teachers needed longer to work with this approach to see results - "*the teachers could've got more out of the project if we could have done it over a year*" but also commented that she felt the teachers had been empowered by the theory. The mentor's view was that this programme would be easier to roll out without rigidly adhering to 2 week cycles. The strength of the project in terms of professional development was seen as the potential of the review sessions as a space and time to reflect. The mentor explained that having 3 teachers in the 45-minute review sessions meant that they had only 15 minutes to talk and she felt this wasn't enough. In her words "potentials have been lost because of time constraints". The mentor would have preferred to have been given a programme structure with the freedom to adapt to fit the school and a longer timeframe to embed.

In terms of scaling up, the mentor felt that currently there was not a great deal of interest from teachers in other year groups, largely because it is a large school with over 100 staff and they tend to focus on their own area. However, she felt that none of the staff would argue with the principles and that in time it would have "naturally disseminated". To supplement the success in dialogic teaching in year 5, the mentor in this school anticipated 'rolling out' training and techniques across the school, probably one

year at a time. One comment made by the teacher in this school ("Ofsted threw us for 3 weeks") suggested that the programme, particularly in its implementation phase, can fall prey to other institutional priorities.

The mentor in School B felt that there was an unrealistic expectation in terms of the amount of suggested extra reading the teachers and mentors could do. She also explained that frequently she was directing the teachers to, for example, the 9 different styles of talk, rather than the teachers bringing this up as part of their reflection and review: "*They're busy class teachers and I have to accept that*." The issue of teachers not being with their class was raised and there was a perception that teachers could be away from teaching too much: "*the classes need the teacher to be teaching*". This raises significant questions for the training and ongoing development that the programme demands in school where teachers are expected to be with their classes as much as possible.

In terms of scaling up the project, the mentor explained how two NQTs had overheard discussion about the project and as a result observed one of the teachers and had shown interest in the laminated 'talk moves' sheet. The intention in the school is for the Year 5 teachers to roll it out to Year 4 and as pupils go up to Year 6 they will continue to practise this approach.

The mentor mentioned "taking out the good stuff" several times without expanding on what this might be, as if the programme itself was too dense: "I would probably just take from it the really good stuff and know that you can change the timings slightly to what suits." The question sheets were seen as particularly useful and the mentor intended rolling these out to the wider staff: "The laminated question sheets are the best bit. We will be encouraging all the staff to use them after the project." The most positive result from this mentor's point of view has been the opportunity to video and review teaching. Furthermore, the class teacher sharing the video with the students "to show children what they're doing" was also seen as a positive.

The mentor in School C provided documentary evidence of the progress made in integrating the dialogic teaching approach with classroom objectives, schemes of work and professional development. These documents detail the careful, staged progression of dialogic teaching activities in the school. School C teachers had evidently benefitted greatly, as attested by the mentor and the teachers, from focussing on discussions of videos of their classes with their mentors. Having the opportunity to both discuss the implementation of dialogic teaching strategies, with evidence from the video, and planning for succeeding lessons was mentioned on a number of occasions as being central to maintaining the momentum of the project and as being of significant value to the teachers and the school. This mentor was confident that the principles and lessons learnt from the project could be cascaded through the school, but recognised that this might take some time. As with school B, the mentor and the teacher commented on the inherent value of reviewing the video lesson had for professional development generally and for this project in particular.

Appendix H: Analysis framework for case studies

Indicators adapted from Alexander (2015 p. 41-44), to provide a framework to guide classroom observations. The original set of indicators included 61 items. From these, the evaluation team removed items that were not likely to be observable during a single lesson observation. Examples of these relate to pupil-pupil talk, and items relating to teachers' reflections on their practice (such as item 14 in the original; 'teachers recognize that in all aspects of classroom talk they themselves are influential models'). Note that the list of indicators is not intended as a checklist, or a measure of the extent to which dialogic teaching has been implemented in schools. It is intended only as a framework to guide the emphasis of classroom observations.

Types of indicators		Individual indicators (with # reference number)	Evidence of indicator found in
A. Contexts & Conditions	I. Dialogic teaching facilitated by:	1. class organisation matches goal	School B
		2. class layout shows flexibility	School B
		3. minimum distractions / interruptions	School B
		4. economical phases limited by a) time required; b) children's attention	School B; School C
		5. intros / conclusions focus on ideas over procedure	
		6. task planning accounts for talking as much as writing	School C
		7. close attention to time-on-task	School B; School C
		8. pace allows efficient coverage of cognitive ground	School C
		9. sustained interaction	School B; School C
		10. increased time for oral tasks	School B; School C
		11. improved assessment of pupils' oral skills	
B. Characteristics	II. Dialogic teaching:	12. questions provoke thoughtful answers	School B
		13. thoughtful answers provoke further discussion	School B
		14. teacher-pupil exchanges are cohesive & extended	School B
		15. balance between talk to participate and talk to explain	School B
		16. everyone asks & explains	School B
		17. turns are cooperative not competitive	School A; School B; School C
		18. non-speakers participate actively	School A
		19. speech is clear, audible, expressive	School B
		20. respond to registers required of different subjects	
		21. mistakes are a chance to learn	
	III. Pupil-pupil interaction:	22. children listen carefully to each other	School A; School B; School C
		23. participate & share ideas	School A; School B; School C
		24. build on others' contributions	School B; School C
		25. work towards common understanding & conclusion	School B; School C
		26. respect minority views	School B
	IV. Teacher-pupil monitoring:	27. lasts long enough to make a difference	School C
		28. emphasises instruction over supervision [scaffolding]	School C
		29. provides useful diagnostic feedback	School A; School B
	V. Questioning:	30. relevant to context and content of lesson	School B; School C
		31. "builds on previous knowledge"	School A; School B
		32. shows understanding	School B; School C
		33. allows for range of responses (open / speculative &c.)	School A; School B
		34. only occasionally predictable	School A
		35. encourages thought & reasoning	School B; School C
		36. provides guidance to avoid errors	
		37. matches form & intent (question=question; instruction=instruction)	School C
		38. allows time to think	School C
	VI. Responses:	39. genuine answers (not looking for 'right' answer)	School B
		40. develop extended answers with types of reasoning	School A; School B
		41. discursive when needed	
	VII. Feedback:	42. informative, useful & diagnostic	School B
		43. reformulates clearly	School C
		44. praises with discrimination	
		45. opens lines of enquiry	School B; School C
		46. encourages open discussion without fear	School B; School C

Appendix I: Data tables for development team video analysis (provided by the project development team)[8]

Talk Moves (English)	Group	N	Mean	Std. Deviation	Std. Error Mean
Teacher open questions	Control	6	10.833	8.010	3.270
	Intervention	15	11.533	5.330	1.376
Teacher closed questions	Control	6	35.333	13.125	5.358
	Intervention	15	35.400	16.690	4.309
Teacher follow-up	Control	6	9.000	6.603	2.696
	Intervention	15	9.933	6.442	1.663
Pupil extended contributions	Control	6	18.833	11.754	4.799
	Intervention	15	21.333	8.226	2.124
Pupil brief contributions	Control	6	35.333	13.125	5.358
	Intervention	15	35.400	16.690	4.309

Appendix I, Table 1: Teacher and pupil talk in English, September 2015

Talk Moves (English)	Group	N	Mean	Std. Deviation	Std. Error Mean
Teacher open questions	Control	9	3.333	3.905	1.302
	Intervention	15	17.733	5.824	1.504
Teacher closed questions	Control	9	35.444	12.001	4.000
	Intervention	15	19.000	5.516	1.424
Teacher follow-up	Control	9	9.000	9.631	3.210
	Intervention	15	15.200	9.756	2.519
Pupil extended contributions	Control	9	10.667	9.014	3.005
	Intervention	15	32.933	12.098	3.124
Pupil brief contributions	Control	9	33.222	8.105	2.702
	Intervention	15	20.333	7.743	1.999

Appendix I, Table 2: Teacher and pupil talk in English in February 2016

[8] These data tables taken from Alexander with Hardman (2017).

Talk Moves (Maths)	Group	N	Mean	Std. Deviation	Std. Error Mean
Teacher open questions	Control	10	7.100	4.909	1.552
	Intervention	15	7.400	7.298	1.884
Teacher closed questions	Control	10	32.100	12.957	4.097
	Intervention	15	41.133	20.546	5.305
Teacher follow-up	Control	10	8.000	8.994	2.844
	Intervention	15	8.733	6.595	1.703
Pupil extended contributions	Control	10	12.800	8.766	2.772
	Intervention	15	16.000	9.979	2.576
Pupil brief contributions	Control	10	29.600	7.291	2.306
	Intervention	15	36.933	16.127	4.164

Appendix I, Table 3: Teacher and pupil talk in mathematics, September 2015

Talk Moves (Maths)	Group	N	Mean	Std. Deviation	Std. Error Mean
Teacher open question	Control	9	1.222	1.641	0.547
	Intervention	15	14.667	5.740	1.482
Teacher closed questions	Control	9	40.222	15.450	5.150
	Intervention	15	19.267	6.995	1.806
Teacher follow-up	Control	9	5.222	3.930	1.310
	Intervention	15	22.667	17.020	4.394
Pupil extended contributions	Control	9	6.222	3.993	1.331
	Intervention	15	35.533	17.691	4.568
Pupil brief contributions	Control	9	35.000	10.642	3.547
	Intervention	15	19.600	6.791	1.753

Appendix I, Table 4: Teacher and pupil talk in mathematics, February 2016

Talk Moves (Science)	PHASE 1	N	Mean	Std. Deviation	Std. Error Mean
Teacher open question	Control	7	8.000	5.033	1.902
	Intervention	10	13.300	6.929	2.191
Teacher closed questions	Control	7	28.286	8.139	3.076
	Intervention	10	31.700	10.874	3.439
Teacher follow-up	Control	7	6.143	5.843	2.209
	Intervention	10	12.200	4.341	1.373
Pupil extended contributions	Control	7	13.000	8.583	3.244
	Intervention	10	25.100	4.508	1.426
Pupil brief contributions	Control	7	28.286	8.139	3.076
	Intervention	10	33.100	8.925	2.822

Appendix I, Table 5: Teacher and pupil talk in science, September 2015

Talk Moves (Science)	PHASE 2	N	Mean	Std. Deviation	Std. Error Mean
Teacher open question	Control	7	4.286	3.094	1.169
	Intervention	10	20.800	7.495	2.370
Teacher closed questions	Control	7	33.286	11.572	4.374
	Intervention	10	20.900	6.855	2.168
Teacher follow-up	Control	7	4.000	2.828	1.069
	Intervention	10	21.500	12.826	4.056
Pupil extended contributions	Control	7	8.143	5.336	2.017
	Intervention	10	42.300	17.994	5.690
Pupil brief contributions	Control	7	31.143	10.205	3.857
	Intervention	10	20.900	6.855	2.168

Appendix I, Table 6: Teacher and pupil talk in science, February 2016

Sub-types of pupil extended contributions	ENGLISH		MATHS		SCIENCE	
	Intervention (10)	Control (8)	Intervention (10)	Control (8)	Intervention (10)	Control (8)
Pupil expand/add	13	4	5	-	7	3
Pupil connect	-	-	1	-	3	-
Pupil explain/analyze	44	24	28	17	52	35
Pupil rephrase	2	4	5	1	1	2
Pupil narrate	2	1	-	-	1	2
Pupil evaluate	5	1	-	-	1	-
Pupil argue	34	4	39	2	19	4
Pupil justify	20	4	35	-	41	4
Pupil speculate	6	3	-	-	-	2
Pupil challenge	3	-	8	-	6	-
Pupil imagine	2	12	2	-	-	1
Pupil shift of position	-	-	3	-	1	-
Total	**131**	**57**	**126**	**20**	**132**	**53**
Mean frequency	**13.1**	**7.12**	**12.6**	**2.5**	**13.2**	**6.62**

Appendix I, Table 7: Comparison of pupil talk in intervention and control groups February 2016

This document is available for download at **www.educationendowmentfoundation.org.uk**

Education
Endowment
Foundation

The Education Endowment Foundation
9th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP
www.educationendowmentfoundation.org.uk