



Education  
Endowment  
Foundation

# Calderdale Excellence Partnership: IPEELL

## Evaluation report and executive summary

November 2018

### Independent evaluators:

Carole J Torgerson, Hannah Ainsworth, Kerry Bell, Louise Elliott,  
Imogen Fountain, Louise Gascoine, Catherine E Hewitt, Adetayo Kasim,  
Dimitra Kokotsaki and David J Torgerson



Durham  
University



UNIVERSITY  
*of York*



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

Education Endowment Foundation  
9th Floor, Millbank Tower  
21–24 Millbank  
SW1P 4QP  
p: 020 7802 1679  
e: [danielle.mason@eefoundation.org.uk](mailto:danielle.mason@eefoundation.org.uk)  
w: [www.educationendowmentfoundation.org.uk](http://www.educationendowmentfoundation.org.uk)

## About the evaluator

The project was independently evaluated by a team from Durham University and the University of York:

Carole J. Torgerson, David J. Torgerson, Hannah Ainsworth, Catherine E. Hewitt, Adetayo Kasim, Dimitra Kokotsaki, Kerry Bell, Louise Elliott, Louise Gascoine,

The lead evaluator was Professor Carole Torgerson.

**Contact details:**

Professor Carole Torgerson  
Professor of Education  
Durham University  
School of Education  
Leazes Road  
Durham  
DH1 1TA

Email: [carole.torgerson@durham.ac.uk](mailto:carole.torgerson@durham.ac.uk)

Tel: 0191 334 8382

## Contents

Executive summary.....	4
Introduction .....	6
Methods .....	11
Impact evaluation .....	20
Implementation and process evaluation .....	38
Conclusion.....	55
References .....	57
Appendix A: EEF cost rating .....	58
Appendix B: Security classification of trial findings.....	59
Appendix C: School Expression of Interest Form.....	63
Appendix D: School Agreement to participate .....	65
Appendix E: Information Letter to Parents/Carers and Opt Out Form .....	69
Appendix F: Supplementary School agreement to participate following changes to the primary outcome .....	71
Appendix G: Supplementary parent/carer letter following changes to the primary outcome .....	74
Appendix H: Trial 2 Primary Outcome Measure .....	76
Appendix I: Sampling framework for school visits & schools visited.....	81
Appendix J: Train the Trainers Observation Schedule.....	84
Appendix K: Train the Teachers Observation Schedule.....	87
Appendix L: School Observation Schedule .....	90
Appendix M: Teacher Interview Schedules.....	93
Appendix N: Pupil Focus Group Schedules .....	94
Appendix O: Sensitivity Analyses of Primary Outcome Data .....	95

---

<b>Appendix P: Subgroup Analyses of Secondary Outcomes.....</b>	<b>102</b>
<b>Appendix Q: Missing Data Analysis for Primary Outcomes.....</b>	<b>105</b>
<b>Appendix R: Distribution of Pre-test Scores .....</b>	<b>109</b>
<b>Appendix S: Additional Results .....</b>	<b>115</b>
<b>Appendix T: Distribution of Post-test scores .....</b>	<b>117</b>

## Executive summary

### The project

The IPEELL intervention is a writing process that encourages students to plan, draft, edit, and revise their writing. IPEELL stands for Introduction, Point, Explain, Ending, Links, and Language. The intervention was developed by the Calderdale Excellence Partnership (CEP) and is adapted from the U.S. programme Self-Regulated Strategy Development (SRSD). In addition to the writing process, the IPEELL intervention also involves ‘memorable experiences’ for students designed to act as a stimulus for their writing. This project tested the impact of one year of IPEELL for children in Year 6 and the impact of the two years of IPEELL children who started it in Year 5 and continued in Year 6. IPEELL was delivered to all children in these year groups.

The project was a randomised controlled trial; 84 schools and 2,682 children participated in the one-year trial and 83 schools and 2,762 children participated in the two-year trial. The trial was an effectiveness trial which tested the intervention under ‘real world’ conditions and schools were randomised to either receive the intervention or to act as a ‘business as usual’ control group. Writing outcomes were measured using Key Stage 2 (KS2) writing outcomes for the one-year trial and a bespoke writing test based on historic KS2 writing tests for the two-year trial. The process evaluation involved observations of the training sessions (both the teacher training and the training of the trainers), observations of lessons, and interviews with teachers. The trial took place between September 2015 and July 2017.

### Key conclusions

1. Pupils who used IPEELL for two years made around two months’ additional progress in writing, compared to pupils who did not. This result has a high security rating. The result was similar when considering only pupils eligible for free school meals who used IPEELL, who also made around two months’ additional progress.
2. The evaluation also measured writing results for pupils who used IPEELL for only one year, using a different writing test. This test showed that after one year, IPEELL pupils had made *less* progress than comparison pupils. This result had a very high security rating. Again, the results were similar for pupils eligible for free school meals.
3. These writing results are not statistically significant. This means that even if the intervention had no impact, the probability that, in this trial, we would have observed an effect size as large as the one found is greater than 5%.
4. Pupils who used IPEELL for two years made less progress in reading, spelling, and maths compared to pupils who did not. It is possible that an increased focus on writing led to less time spent on these topics.
5. The training of teachers was lacking in both the provision and breadth of practical examples and in fully modelling some aspects of the approach. This may have been because the training in this trial was carried out by inexperienced trainers and it may, therefore, be useful for the trainers to watch an experienced trainer train the teachers first.

### EEF security rating

The findings for all pupils in the one-year trial has very high security and the finding for the two-year trial has high security. The findings for FSM pupils have lower security ratings – for the one year trial the FSM outcome has high security and for the two year trial the FSM outcome has moderate to high security. The outcomes were measured through a large effectiveness trial, which was a well-designed randomised controlled trial and the pupils in the IPEELL schools were similar to those in the comparison classes in terms of prior attainment. However, for the two-year trial the security was reduced due to 20% of all of the pupils who started the trial not being included in the final analysis as post-test data

were not collected for these pupils. For the FSM outcomes, the one year trial was not as well powered due to a smaller sample of FSM pupils, whilst for the two year trial over 20% of the pupils who started the trial were not included in the final analysis due to post-test data not being collected for these pupils.

### Additional findings

The one-year and two-year trials of IPEELL show different impacts on writing. The one-year trial suggests a negative impact, while the two-year trial suggests a positive impact. One reason for this difference could be that different writing outcomes were used for the two trials, although it does not help to explain the apparent difference in the direction of the effect. During the set-up of this trial, the way KS2 writing outcomes were recorded was changed. Rather than being given a score, pupils were allocated to one of three categories based on the extent to which they met expected levels. Using the new measure could reduce the likelihood of the trial identifying an impact as it would only register improvements where children moved between categories and would not show if children improved within a category. Because of this, the outcome for the two-year trial was changed to a bespoke writing test based on historical questions from the old KS2 writing tests, allowing for the impact of the programme on a continuous outcome to be recorded. The one-year trial therefore shows impacts on a categorical outcome, whereas the two-year trial shows impacts on a continuous outcome.

A previous EEF-funded trial of IPEELL in England showed positive impacts. In contrast, this project did not find evidence of consistent positive effects. This could be because this trial had several key differences to the previous trial, including: the teacher training and implementation monitoring being delivered by newly recruited trainers instead of the developers of IPEELL; the intervention was delivered in Years 5 and 6 rather than Years 6 and 7, meaning that it was not used to support pupils through the transition to secondary school; and the intervention was used with children across the full range of prior attainment rather than with low prior attainers only. Interestingly, for the two-year trial, we do see a more positive result for low prior attainers of three months' additional progress, although this result relates to a smaller group of children so will have lower security than the overall result.

As well as the impact on writing attainment, the trial assessed whether IPEELL has 'spillover' effects on maths, reading, and spelling. In the two-year trial, IPEELL pupils made significantly less progress in reading, maths, and spelling than comparison pupils who did not use IPEELL. The results were similar for the one-year trial, although in that case only the reading result was significant. It is not clear from the evaluation why this happened, however one hypothesis is that class time was directed away from these subjects and towards the teaching of writing.

### Cost

The cost of the approach is estimated at £36 per pupil per year when averaged over three years. Teachers delivering the intervention have to attend one day of training and schools would need to meet the cost of cover for this.

**Table 1: Summary of impact on primary outcome**

Group	Effect size (95% confidence interval)	Estimated months' progress	EEF security rating	No. pupils	EEF cost rating
2 years of IPEELL – impact on writing	0.11 (-0.13, 0.34)	2	🔒🔒🔒🔒🔒	2196	£££££
2 years of IPEELL (FSM) – impact on writing	0.14 (-0.14, 0.41)	2	🔒🔒🔒🔒🔒	811	£££££
1 year of IPEELL – impact on writing	-0.09 (-0.34, 0.15)	-1	🔒🔒🔒🔒🔒	2465	£££££
1 year of IPEELL (FSM) – impact on writing	-0.07 (-0.38, 0.23)	-1	🔒🔒🔒🔒🔒	962	£££££

## Introduction

### Background evidence

In 2013, approximately 85,000 pupils left primary school without having attained level 4 or above in writing (Department for Education, 2013a). A substantial minority of pupils who do not reach level 4 in English at the end of Key Stage 2 (incorporating reading and writing) do not achieve the expected level of progress in secondary school (Department for Education, 2014).

The IPEELL intervention is a writing process model in which students are encouraged to plan, draft, edit, and revise their writing. IPEELL stands for Introduction, Point, Explain, Ending, Links, and Language. The strategy, originating as Self-regulated Strategy Development (SRSD) in the United States in the 1990s, provides a clear structure to assist writers and can be used for most genres of writing, including narrative writing. There are six basic stages of instruction and four strategies for self-regulation, which include self-monitoring and goal setting, thus providing pupils with ownership for improving their own writing. IPEELL uses 'heuristics' which provide scaffolding of structures and devices that aid the composition of argumentative writing—in particular planning—which can include examining a question, brainstorming, organising and sequencing ideas, and evaluating. The approach was designed, and is suitable, for children who are aged between 8 and 14 (Andrews et al., 2006).

A systematic review conducted by Andrews et al. (2006) found that a combination of contextual factors and specific interventions based on the SRSD programme were necessary for successful practice in teaching and learning writing for 7–14-year-olds. These included interventions based on a writing process model (planning, drafting, editing, and revising) which encouraged self-motivation (personal target setting). The effect sizes of individual studies, largely undertaken in the U.S., were very large with estimates in some instances exceeding one standard deviation between the intervention and control groups. For example, in a study by De La Paz and Graham (2002) of secondary school children of mixed attainment (aged between 12–14) the overall quality of their writing was 1.7 standard deviations better than the quality of writing of children in the control group six weeks after an intervention to develop planning activities such as organising and sequencing ideas. Nevertheless, despite these promising results in a U.S. context it was important to establish whether or not such an intervention would be effective in a U.K. setting.

In 2013/2014, an efficacy trial undertaken by some of the authors amongst Year 6 and 7 pupils (who had not achieved a level 4 at the beginning of Year 6) in 23 British schools of a self-regulated strategy development instruction intervention combined with 'memorable experiences' showed a large improvement compared with control children not exposed to the intervention (effect size 0.75; Torgerson et al., 2014). This large difference was in line with previous, mainly United States-based, studies of self-regulated strategy development alone.

The present effectiveness trial aimed to evaluate the impact of the intervention in a trial of 'roll-out' before widespread implementation. Additionally, the evaluation sought to evaluate a 'train the trainers' approach to delivery (described in more detail below), as well as a number of other important issues, including generalisability, 'Anglicisation' of the programme, and outcome measurement.

Firstly, the issue of generalizability is partially addressed. The schools in the efficacy trial were from a small area of West Yorkshire and they had worked together for a number of years; it is possible that these results may not be fully generalizable to the wider population. This issue was addressed by recruiting schools from two different geographical locations: Leeds and Lincolnshire. Characteristics of participating schools compared to national norms (Table 2) indicate how representative they are of the national picture. About 8% and 74% of the recruited schools were rated as 'outstanding' and 'good' in comparison with an English average of 19% and 72%, respectively. The average size of the recruited



schools was similar to an English average of 260 pupils per school. However, the percentage of children eligible for FSM (free school meals) in the recruited schools is twice the English percentage of 14.5%.

**Table 2: Comparison of school characteristics and England average**

Factor		IPEELL	England
Ofsted Rating	Outstanding	8%	19%
	Good	74%	72%
School Size		253	260
FSM		33.4%	14.5

Secondly, in the delivery of the previous efficacy study, the intervention was adapted or ‘anglicised’ and this modified version warranted an effectiveness evaluation. The modified, anglicised version of the intervention was used in this trial.

Thirdly, the main outcome measure used in the previous trial, Progress in English 6 (PiE6), is publically available to teachers and therefore, in theory, it would have been possible for teachers to have had sight of the test and have been influenced by the test before it was delivered to the pupils as post-test. The test also provides children with different scaffolding instructions for planning their writing for the test which are in contrast to the scaffolding advocated by IPEELL. In this trial we originally planned to use KS2 results as the primary outcome for both Trial 1 and Trial 2. However, due to changes in the way that KS2 assessments are recorded in the National Pupil Database which came into effect in the academic year 2016/2017, KS2 results alone were deemed insufficient to fully assess the effectiveness of the intervention. (A more detailed discussion of this can be found in section Primary Analysis.) Consequently, the primary outcome in Trial 2 was an independently administered and marked test that drew on components of historical KS2 assessments. The published protocol was updated on 22 March 2017 to reflect this change.

## Intervention

The Education Endowment Foundation commissioned Durham University and the University of York to evaluate IPEELL, a literacy-based intervention centred on self-regulated strategy development being delivered by Calderdale Excellence Partnership between 2015 and 2017.

The intervention is a continuing professional development (CPD) training programme for Year 5 and Year 6 teachers which covers IPEELL as well as the facilitation of ‘memorable experiences’. In this report, these two aspects are collectively referred to as ‘IPEELL’.

In addition to providing a structure to assist with writing, the intervention also emphasises the value of memorable experiences for writing. It is a widely held belief that children being raised in low income households may have little access to novel environments such as day trips to the coast or places of historical interest. For such children, being provided access to such experiences through school can provide memorable and stimulating experiences that can provide a focus for learning. The intervention suggests a wide range of positive writing possibilities that can accrue from memorable experiences and encourages schools to facilitate such endeavours. As part of the evaluation, schools received funding to ensure participating pupils engage in ‘memorable experiences’ such as school trips or unusual

activities taking place in school. The purpose of these is to provide stimuli for children's writing practice during the year.

A pragmatic approach was taken to training with a 'train the trainers' cascade model being implemented. This differed from the training model evaluated in the efficacy trial where the Calderdale Excellence Partnership delivered the training directly to all schools. Members from the Calderdale Excellence Partnership conducted training sessions for trainers from Leeds Local Authority and the Lincolnshire Centre for British Teachers (CfBT). These trainers then delivered the CPD training programme to teachers in their respective areas. Each teacher participated in a day long training where the project was presented and various group activities helped consolidate different aspects of IPEELL. The training was delivered twice in each of the two areas (Leeds and Lincolnshire) but to two different groups of teachers. Following the training day, the teachers had the opportunity to take part in networking meetings where they could share ideas around implementation with other teachers. Schools received a teachers' manual prepared by the Calderdale Excellence Partnership and support was on hand where required from the team. Questions regarding the use of the manual were explored in the focus group interviews with the teachers and are discussed below (see 'Implementation and Process Evaluation'). Teachers also had the opportunity to ask questions and clarify issues around the use of the manual when members from the Calderdale Excellence Partnership visited their school. The implementation team quality assured the training programme throughout. Information about the training, including formative findings from our observations, is discussed in the Outcomes section of the Process Evaluation section of the report.

## Evaluation objectives

### Aim of impact evaluation

To estimate the impact of IPEELL delivered to pupils in Years 5 and 6 on writing outcomes when rolled out at scale.

### Impact evaluation primary research questions

Trial 1: What is the effectiveness of IPEELL when delivered for one year to pupils in Year 6 compared with 'teaching as usual' on the writing skills of participating children?

Trial 2: What is the effectiveness of IPEELL when delivered for two years to pupils beginning in Year 5 and moving into Year 6 compared with pupils receiving 'usual teaching' on the writing skills of participating children?

### Impact evaluation secondary research questions

What is the impact of IPEELL on children ever eligible for Free School Meals (FSM)?

What is the impact of IPEELL on academically high achievers and academically low achievers?

### Aim of process evaluation

The process evaluation of the writing intervention had two main aims:

1. to evaluate the fidelity of implementation of the intervention as described by the Calderdale Excellence Partnership; and
2. to investigate issues of implementation and organisation.

## Ethics and trial registration

Ethical approval was obtained from Durham University School of Education (SoE) Ethics Committee on 9 February 2015 and via Chair's action from the Health Sciences Research Governance Committee at the University of York on 30 June 2015. The protocol was revised from the original design to account for a change to the primary outcome measure. The revised protocol was approved by Durham University SoE Ethics Committee on 1 March 2017 and by Chair's action from the Health Sciences Research Governance Committee at the University of York on 22 March 2017.

All study documents including those used in the process evaluation were reviewed as part of the ethical review.

The trial is registered with ISRCTN, reference number ISRCTN11002306.

## Data protection

All data were stored and processed in accordance with the Data Protection Act (1998).

Schools were informed of the data requirements at recruitment events and through the memorandum of understanding. All parents/carers of pupils in the trial classes received an information sheet, see Appendix E, which outlined the data schools were providing about the pupils in the trial and how it would be used. Parents/carers were given the option to withdraw their child from data sharing.

Schools provided pupil details (name, unique pupil number (UPN), date of birth, gender, FSM status, and Pupil Premium status for all pupils in the trial classes at baseline to allow the evaluation team to request KS1/KS2 results and check FSM status for these pupils from the National Pupil Database and to allow arrangements to be made for independent testing. Access to pupil details was limited to necessary members of the evaluation team, project team, and the organisation conducting independent testing (NFER). Collected data was used for statistical analysis and will be shared with the Department for Education, the Education Endowment Foundation (EEF), FFT Education, and, in an anonymised form, to the U.K. Data Archive.

All results will be anonymised so that no school or individual pupil will be identifiable in the report or dissemination of any results.

## Project team

The intervention was delivered by the project team—Calderdale Excellence Partnership—comprising:

Ken Inwood  
Carlton Midgley  
Amanda Rawson

The project team was responsible for overseeing all aspects of the training model, monitoring in schools, and providing guidance to teachers.

## Evaluation Team

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

Durham University: Carole Torgerson, Louise Gascoine, Adetayo Kasim, Dimitra Kokotsaki.

University of York: Hannah Ainsworth, Kerry Bell, Louise Elliott, Imogen Fountain, Catherine Hewitt, and David Torgerson.

Carole Torgerson was the responsible grant holder and provided oversight to all aspects of the evaluation.

Independent testing in Trial 2 was conducted by NFER.

## Methods

### Trial design

<b>Trial type and number of arms</b>	<b>Cluster randomised trial with two parallel arms</b>
<b>Unit of randomisation</b>	School
<b>Stratification variable(s) (if applicable)</b>	Number of pupils on roll, proportion of children currently eligible for free school meals (FSM), geographical area, and mixed year group teaching (of Year 5 and 6 pupils).
<b>Primary outcome</b>	<b>variable</b> Trial 1: KS2 writing results at one year follow-up. Trial 2: A study-specific writing test that drew on components of historical KS2 assessments.
	<b>measure (instrument, scale)</b> Trial 1: Ordinal scale. Trial 2: Continuous scale.
<b>Secondary outcome(s)</b>	<b>variable(s)</b> Historical KS2 writing test (only for Trial 2), KS2 reading (national test), KS2 spelling, punctuation and grammar (national test) and KS2 maths results (national test).
	<b>measure(s) (instrument, scale)</b> Historical KS2 Writing test (only for Trial 2): Ordinal scale. Others: Continuous scale.

Two pragmatic cluster randomised controlled trials were combined within the same evaluation.

Trial 1 evaluated the impact of pupils receiving IPEELL for one year in Year 6 compared with usual teaching on the high stakes and policy relevant KS2 results.

Trial 2 evaluated the impact of pupils receiving IPEELL for two years, Years 5 and 6, compared with usual teaching on both the high stakes and policy-relevant KS2 results and an independently-administered and marked test that drew on components of historical KS2 assessments.

All children will be followed up longitudinally beyond the trial period through the National Pupil Database (NPD).

Eighty four schools across Lincolnshire and Leeds area were recruited and randomised on a 1:1 basis. Schools in the intervention group received training in IPEELL from the project (implementation) team during June 2015. Schools in the intervention group were supported to deliver IPEELL to Year 5 and 6 pupils during the academic year 2015/2016 beginning in September 2015 and to Year 6 pupils in the 2016/2017 academic year. Schools in the control group were asked to continue with usual teaching with Year 5 and 6 pupils in the 2015/2016 academic year. From June 2016, control schools were able to receive training and support to deliver IPEELL to pupils in Year 5 (and Year 4 if they wished) in the

2016/2017 academic year. Schools were explicitly asked not to deliver IPEELL to pupils in Year 6 in the 2016/2017 academic year as these constituted the control children in the Trial 2 analysis. As a consequence of this, there was some potential for contamination or spillover effects in the Trial 2 comparison, for instance, if schools in the control group whose teaching staff had received training chose to deliver IPEELL or elements of IPEELL to Year 6 children in the 2016/2017 academic year. The process evaluation sought to determine the extent to which this may have been the case. A cluster randomised design was selected to minimise contamination effects between intervention and control participants. A financial incentive was offered to all schools to aid with retention to the end of the research.

The trial was designed, conducted, and reported following the CONSORT standards (Altman et al., 2011) in order to minimise all potential threats to internal validity such as selection bias and a range of post-randomisation biases (Cook and Campbell, 1969; Shadish, Cook and Campbell, 2002; Torgerson and Torgerson, 2008).

## Participant selection

### Recruitment

CEP and the evaluation team worked collaboratively in the recruitment of schools. CEP was primarily responsible for recruitment whilst the evaluation team provided documentation on the trial for the schools and pupils/parents including details of the trial design.

The recruitment strategy preferentially targeted schools with high proportions of FSM pupils and high proportions of children achieving below age-expected levels at KS2. In each of the geographical areas (Leeds and Lincolnshire), CEP initially held information events at which schools were invited to complete an expression of interest form. Following these initial meetings, four further recruitment events were held (two in each area) by CEP and the evaluation team, the purpose of which was to inform schools about the project (including information about the intervention, pupil eligibility criteria, data requirements, and design of the evaluation) and to invite them to sign an Agreement to Participate Form (Appendix D). A number of techniques were used to contact schools and invite them to the events including postal invitation, direct email (where possible to the headteacher or alternatively a general school email address), websites, and headteachers' meetings.

All participating schools were offered £750 to compensate them for the extra time needed to provide the data to the evaluators for analysis. It was anticipated that this payment would also ensure that the control schools were less likely to drop-out after randomisation.

### Schools

Primary schools (excluding those in Ofsted category four, 'Inadequate') were recruited from the Leeds and Lincolnshire areas. Schools were asked to sign an Agreement to Participate Form (Appendix D) to demonstrate their commitment to the project and understanding of the evaluation requirements. Any schools which were allocated to the control group and which were unable to ensure that pupils in Year 6 in the 2016/2017 academic year would not receive IPEELL (for example schools which had mixed Year 5 and Year 6 classes) were excluded from the Trial 2 analysis but remained in the study for Trial 1. Following the changes to the trial design relating to the primary outcome measure, schools were requested to sign an updated Agreement to Participate (Appendix F) to confirm that they would facilitate the additional assessment.

### Pupils

All Year 5 and Year 6 pupils in the 2015/2016 academic year at participating primary schools were eligible for inclusion in the study. Pupils were identified from the school roll in the 2014/2015 academic

year (when they were in Years 4 and 5). Schools informed parents of all pupils in Years 4 and 5 by means of a letter about the study (material and help provided by the evaluation team and Calderdale Excellence Partnership—Appendix E). Parents were given the opportunity to withdraw their child's data from being used in the evaluation (opt-out consent). Pupils whose parents requested opt-out of their child's data for the purposes of the research received the intervention, as this was delivered to the whole year, but their data were not shared.

Following the changes to the trial design relating to the primary outcome measure, schools sent an updated parent information letter and opt-out consent form to parents (Appendix G) to confirm that they did not object to their child's data being shared with an additional party.

### **Teachers**

There were no selection criteria for teachers. All teachers teaching Year 5 (2015/2016) and Year 6 (2016/2017) were eligible.

### **Outcomes measures**

#### **Baseline data**

Schools were asked to provide the name, unique pupil number, date of birth, and gender for all eligible pupils at baseline to enable linkage to the National Pupil Database (NPD). Baseline attainments were measured using KS1 results described by KS1\_WRITPOINTS, KS1\_READPOINTS and KS1\_MATPOINTS in the NPD for writing, reading, and maths, respectively.

#### **Primary outcomes**

For Trial 1, KS2 writing results (KS2\_WRITTAOUTCOME) at one year follow-up were obtained from the NPD. This is a national standardised test that is assessed by teachers. A potential limitation of the KS2 test is that it is administered and marked by teachers and therefore it is not independent of the delivery of the intervention.

For Trial 2, due to changes in the way that KS2 assessments are recorded in the NPD that came into effect in the academic year 2016/2017, KS2 results alone were deemed insufficient to fully assess the effectiveness of the intervention. The new KS2 assessment asks teachers to define children's progress in writing in terms of 'working towards the expected standard', 'working at the expected standard', and 'working at greater depth within the expected standard' (Standards and Testing Agency, 2016). The move to a categorical approach to assessment may mean that a small difference between the intervention and control group could be missed. A test that drew on components of historical KS2 assessments (past KS2 writing papers) was independently administered and marked (Appendix H) and provided numerical data for analysis. One long persuasive writing task and one short recount writing task were completed. The test was delivered by administrators (blinded to group allocation) from the National Foundation for Educational Research (NFER). NFER was also responsible for blind marking the test and linking the results with pupil details. The published protocol was updated on 22 March 2017 to reflect this change. Histograms of the post-test are presented in Appendix T.

#### **Secondary outcomes**

*Secondary outcomes for Trial 1 included:*

- KS2 reading (national test, KS2\_READSCORE);
- KS2 spelling, punctuation and grammar (national test, KS2\_GPSSCORE); and
- KS2 maths results (national test, KS2\_MATSCORE).

*Secondary outcomes for Trial 2 included:*

- KS2 writing (national test, KS2\_WRITTAOUTCOME);
- KS2 reading (national test, KS2\_READSCORE);
- KS2 spelling, punctuation and grammar (national test, KS2\_GPSSCORE); and
- KS2 maths results (national test, KS2\_MATSCORE)

### *Implementation fidelity*

A tool to record implementation fidelity was developed in collaboration with the implementation team (Appendix L). CEP and the evaluation team used the agreed observation schedule during the school visits. This allowed a fidelity score to be recorded for every participating intervention school that was visited. The schedule was also used when observing control schools (from June 2016) to consider contamination. This gave more robustness to the process evaluation and allowed us to see in a pragmatic way the full range of delivery. A random selection of control schools stratified by key factors was made for these visits (see Appendix I).

From June 2016, control schools were able to receive training and support to deliver IPEELL to pupils in Year 5 (and Year 4 if they wished) in the 2016/2017 academic year. Schools were explicitly asked not to deliver IPEELL to pupils in Year 6 in the 2016/2017 academic year as these constituted the control children in the Trial 2 analysis. As a consequence of this, there was some potential for contamination or spillover effects in the Trial 2 comparison, for instance, if schools in the control group whose teaching staff had received training chose to deliver IPEELL or elements of IPEELL to Year 6 children in the 2016/2017 academic year. The process evaluation sought to determine the extent to which this may have been the case.

A common observation schedule was also used for the 'train the teachers' observation by both the evaluation team and CEP (Appendix K; see section on Methods for the process evaluation below).

### **Sample size**

The sample size calculation for Trial 1 assumed recruitment of 80 schools with 3,600 children in total (45 per school) and an intra-cluster correlation of 0.08 (taken from a previous evaluation of Year 6 pupils; Torgerson et al., 2014) and a correlation with KS1 scores of 0.7 (Torgerson et al., 2014). Based on this we estimated that the sample size would detect a minimum difference of 0.14 standard deviations (80% power,  $2p = 0.05$ ) between the intervention and the control groups.

To allow for attrition, the sample size calculation for Trial 2 assumed a smaller number of 70 schools (excluding schools that could not ensure separate Year 6 teaching) with 3,150 children in total (45 per school). As in Trial 1, an intra-cluster correlation of 0.08 (taken from a previous evaluation of Year 6 pupils; Torgerson et al., 2014) and a correlation with KS1 scores of 0.7 (Torgerson et al., 2014) were assumed. Based on this we estimated that the sample size would detect a minimum difference of 0.15 standard deviations (80% power,  $2p = 0.05$ ) between the intervention and the control groups.

The effect of the intervention was also analysed in the subgroup of pupils who were eligible for FSM. In a previous trial of SRSD (Torgerson et al., 2014), approximately 28% of children were eligible for FSM. Estimating that there will be approximately 750 EVERFSM children in total in Trial 1 with 80 schools (nine pupils per school on average) would allow us to show a minimum difference of an effect size of around 0.23 assuming 80% power, intra-cluster correlation of 0.08, a correlation of 0.7 with KS1 scores, and two-sided test at 5% significant level. Estimating that there will be approximately 750 EVERFSM children in total in Trial 2 with 70 schools (11 pupils per school on average) would allow us to show a minimum difference of an effect size of around 0.22 assuming 80% power, intra-cluster correlation of 0.08, a correlation of 0.7 with KS 1 scores, and two-sided test at 5% significant level. Sample size was



based on a cluster randomised trial with person-level outcomes using Optimal Design Plus Empirical Evidence software.

## Randomisation

Schools were allocated on a 1:1 basis to either receive the intervention for the Year 5 cohort in 2015/2016 (the intervention group) or to receive the intervention for the Year 5 cohort in 2016/2017 (the waitlist school control group). The allocation was undertaken by a statistician in the evaluation team (AK) who was independent of delivery of tests and intervention using anonymised datasets.

Minimisation, a form of random allocation (Torgerson and Torgerson, 2008), was undertaken to ensure the schools were balanced. This approach was taken over simple or stratified randomisation using blocks as this allows for a better balance in terms of observable school-level characteristics compared with these other methods. Minimisation was conducted using the following factors:

- number of pupils on roll;
- proportion of children currently eligible for free school meals (FSM);
- geographical area; and
- mixed year group teaching (of Year 5 and 6 pupils).

These factors were chosen as a proxy for schools' characteristics that could impact performance. For example, pupils on FSM are less likely to perform as well as other children and FSM is also strongly associated with area level deprivation. This means that schools with higher proportions of FSM pupils are likely to be located in deprived areas. While pupil-level prior attainment data like KS1 scores could be considered more appropriate, these data were not available pre-randomisation.

Data relating to these school-level factors and were collected from schools via the delivery team or from publically available information. Recruited schools were randomised in one batch using *MinimPy* software. One school from Leeds had mixed Year 5 and Year 4 classes, but not for Year 6 and Year 5 classes. For this trial, they were therefore treated as having no mixed classes. Another school with mixed classes was treated as having no mixed classes because children are taught in 'age' grouping for literacy. Two schools from Lincolnshire with mixed classes had no clear information on whether they had mixed classes or not for the Year 6 and Year 5, which were treated as having mixed classes in this trial. The percentages of children on FSM (%FSM) were categorised as 'low' or 'high' based on a median value within each local authority (LA). Schools with %FSM less than their LA median were categorised as 'low' and those with %FSM greater than their LA median were categorised as 'high'. Number on roll (NOR) was categorised as low or high within each LA in a similar way as the %FSM.

## Statistical analysis

Analysis was conducted using the principles of 'intention to treat', meaning that all schools and pupils were analysed in the group to which they were randomised irrespective of whether or not they ultimately received the intervention. Statistical significance was assessed at the 5% level unless otherwise stated; 95% confidence intervals are provided as appropriate.

## Primary analysis

The primary analysis used two primary outcomes: KS2 writing scores between the two groups of schools for Year 6 pupils in Trial 1 for those who received one year of IPEELL teaching; and a writing test (based on historical KS2 assessments) for Year 6 pupils in Trial 2 for those who received two years of IPEELL teaching. The primary analysis of KS2 writing scores in Trial 1 was done using a proportional odds model to account for the ordinal nature of the data. The proportional odds model compares the probability of being at one level of writing (or lower) versus being at levels above it. An example is

comparing the probability of writing outcome 'below expected standard' versus 'working at the expected standard or higher'. The estimated odds ratio for an improvement from lower level(s) to higher level(s) of writing outcome between the intervention groups was converted to an effect size as recommended by Borenstein *et al.*, (2009) by using the expression, 'log odds ratio' \*  $\sqrt{3}/\pi$ . Note that Cohen's *d* effect size and Hedges' *g* effect size are equivalent because the multiplicative correction factor in Hedges' *g* effect size is 1 due to having a large sample size. The primary analysis of the outcome in Trial 2 was analysed using a multilevel model with adjustment for minimisation factors in accordance with the updated EEF analysis guidance. Analysis of primary outcome with the KS1 score and the intervention groups as the only covariate is reported as sensitivity analysis to check the impact of the minimisation factors on effect size. Clustering of pupils within schools was accounted for as random effects to obtain robust standard errors and total variance, that is, the sum of within- and between-schools variances (Aerts *et al.*, 2002). The multilevel model was formulated as a random intercept model as defined by Verbeke and Molenberghs (2000). Effect size estimation for Trial 2 accounted for cluster randomisation as proposed by Hedges (2007). Note that the denominator was based on the total variance from multilevel model after adjusting for the factors used in the minimisation scheme.

### Secondary analyses

Secondary outcomes were analysed in the same way as the primary outcomes depending on whether the data were ordinal or continuous in nature. Multilevel models, adjusting for clustering, were used to compare differences between the intervention and control groups with respect to the secondary outcomes of KS2 reading, KS2 spelling, punctuation, and grammar, and KS2 maths results to assess whether there is any evidence of a 'spillover' effect of IPEEL teaching. The corresponding KS1 scores were used as baseline data except for KS2 spelling, punctuation, and grammar where KS1 reading was also used as baseline data because of lack of KS1 spelling, punctuation, and grammar in the merged National Pupil Database. For Trial 2, KS2 writing was also explored as a secondary outcome.

### Subgroup analysis

Subgroup analysis of the 'ever eligible for FSM' group was pre-specified in the trial protocol and was done using a proportional odds model for analysis of the ordinal writing outcome in Trial 1. A multilevel model was used for subgroup analysis for FSM in Trial 2. The effect of the intervention on academically high achievers and academically low achievers was also investigated. Low achievers were defined as those with scores below the national average of 2b in KS1 writing; and high achievers were those with scores of 2b and above in KS1 writing.

### Sensitivity analysis

The primary writing outcome for Trial 1 was also analysed using a multilevel model to investigate whether the effect size from the proportional odds model is very different than if the ordinal data are treated as scores and analysed as a continuous outcome. Analyses based on non-parametric bootstrap confidence intervals and permutation tests were performed to sensitise the estimated effect size in Trial 2. Further sensitivity analyses for missing data were performed at pupil level instead of school level. Cross-tabulation was used to examine the distribution of number of pupils without post-test scores by intervention groups, separately for Trial 1 and Trial 2. A logistic regression drop out model was fitted to identify minimisation factors that were predictive of missing data. Multiple imputation (ten independent imputations) was performed to explore whether missingness mechanisms were sufficiently dependent on the observed data, such as KS1 scores.

## Implementation and process evaluation

### Aims

The process evaluation had two main aims:

1. to evaluate the fidelity of implementation of the intervention as described by the Calderdale Excellence Partnership; and
2. to investigate issues of implementation and organisation.

## Methods

### *Design*

The design was descriptive. It used methods that enabled data to be collected that could describe the extent to which the intervention was implemented with fidelity and also stakeholders' perceptions of implementation and organisation at one moment in time. This 'moment' was the 18 months in which the intervention developed by CEP was delivered using the 'train the trainers' model. The trainers were trained and all Year 5 and Year 6 teachers attended one training day. During this time, the following data collection activities took place.

### *Data collection*

In order to investigate the fidelity of implementation of the intervention, the evaluation team made a number of observations at each stage of delivery, namely the training of the trainers session (first level of training where the trainers were trained by the CEP developers) and then a number of 'train the teachers' sessions (Year 5 and Year 6 teachers were trained together in one-day session). CEP also attended the 'train the teachers' sessions. A common observation schedule was used for the 'train the teachers' observation by both the evaluation team and CEP (Appendix K). The focus was on how well the trainers could replicate the training and how well the teachers could implement the intervention following training in order to investigate how well the training was cascading down to the participating schools and whether the 'training the trainers' model could work.

Stakeholder perceptions of implementation and organisational issues were explored through in-depth process work. The evaluation team carried out twenty formal observation site visits through a random selection of schools stratified by key factors (see Appendix I; note that one of the schools withdrew from the intervention before the end of the project). Each of the site visits included interviews with the teaching staff and focus groups of up to four pupils (to understand the perceptions and experiences of key stakeholders) and also an observation schedule was developed to look at how the programme was being implemented (Appendix L).

In addition to the evaluator process work, CEP visited:

- all intervention schools in the Autumn Term 2015 with the area trainer;
- all or some intervention schools in the Spring and Summer terms of 2016 with the area trainer; and
- all control schools in the 2016/2017 academic year with the area trainer.

The evaluation team and CEP aimed to have a 'joined-up' approach in terms of communication with schools and tried to ensure whenever possible that visits from both parties were made at the same time. Any observations from CEP were checked as a quality assurance measure with observations made by the evaluation team for the process evaluation.

### *Observations*

An agreed observation schedule was used by both CEP and the evaluation team during the school visits (Appendix L). Both parties agreed that there were certain features of the writing project that were non-negotiable which were expected to be seen in every lesson (for example, about the genre, use of mnemonics, the planning frame, positive self-talk, and so on), whereas evidence for other features (such as self-scoring, peer scoring, recording scores on a recording sheet for ease of reference, and do on) could be seen in sources such as wall displays and pupil exercise books. A particular issue we needed to consider was the need to differentiate between measuring fidelity of delivery, which was our main aim, and assessing the quality of the teaching.

### *Interviews*

The interviews with the teaching staff and the pupils aimed to:

- assess the necessary conditions for successful delivery;
- assess the actual and potential barriers to successful delivery;
- identify the perceived outcomes of the intervention, including interest and enjoyment by the pupils; and
- identify any potential negative effects on the pupils.

Interview schedules are provided in Appendix M and Appendix N.

The process evaluation also sought to determine the extent to which teaching staff in the control group who would have received training managed to refrain from delivering elements of IPEELL to Year 6 children in the 2016/2017 academic year (as there was the potential for contamination or spillover effects in the Trial 2 comparison).

We attempted to understand this by initially having a discussion with the Year 6 teachers who assured us that no aspects of IPEELL formed part of their teaching. This was ascertained by inspecting some of the Year 6 pupils' writing workbooks where there was no evidence of the use of the IPEELL writing structure or its different stages.

## **Costs**

### **Intervention costs**

Data on costs were collected directly from the Calderdale Excellence Partnership and included training costs, materials costs, and an estimate of the costs associated with engaging children in memorable experiences.

### **Time costs**

Given that the intervention is designed to be integrated into standard teaching practices, the only additional time required for schools is the time required for training of the teachers. There may also be some additional time required to schedule the memorable experiences but it is not possible to predict how much time this would have taken as it will have varied by school and the nature of the experience. Therefore, no data on time were collected.

## Timeline

**Table 3: Timeline**

Date	Activity
February 2015	Apply for ethical approval
	Recruitment of schools
	Training
September 2015–April 2016	Baseline data collection
March 2015	Randomisation
September 2015–July 2017	Delivery of intervention
Trial 1: May 2016 Trial 2: June/July 2017	Post intervention testing
October 2017	Analysis and report writing of Trial 1
May 2018	Analysis and report writing of Trial 2

## Impact evaluation

### Participant flow including losses and exclusions

Eighty-four schools were randomised.

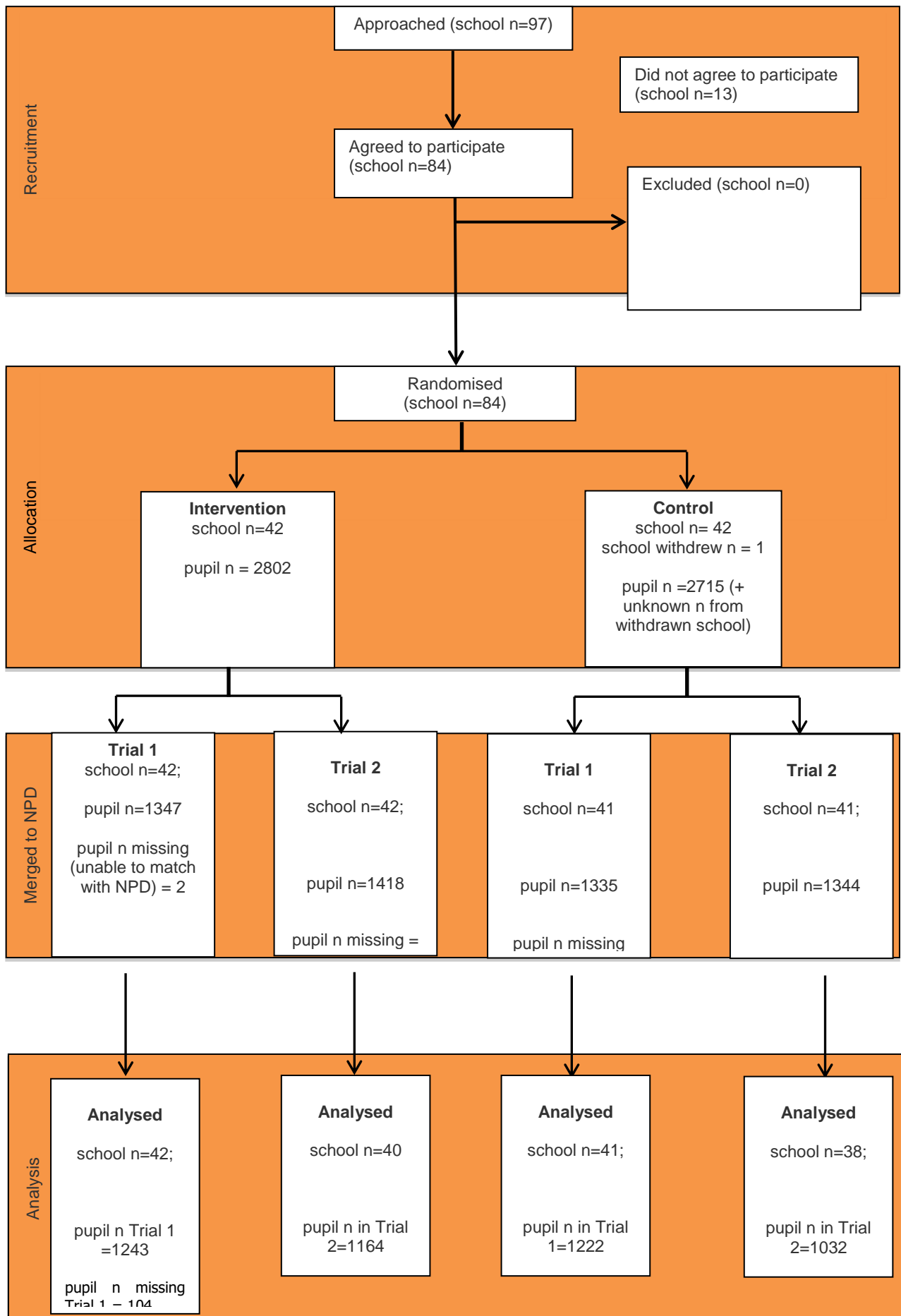
One school allocated to the control group dropped out after being informed of its allocation but before baseline data were collected. As pupil data was not planned to be collected until after randomisation, it was not possible to collect pupil numbers and characteristics from this school.

However, calculations of median number of pupils per school per trial suggest that the median number of pupils for schools was 29 in Trial 1 and 30 in Trial 2. The figure for FSM pupils per school was 10 in Trial 1 and 11 in Trial 2.

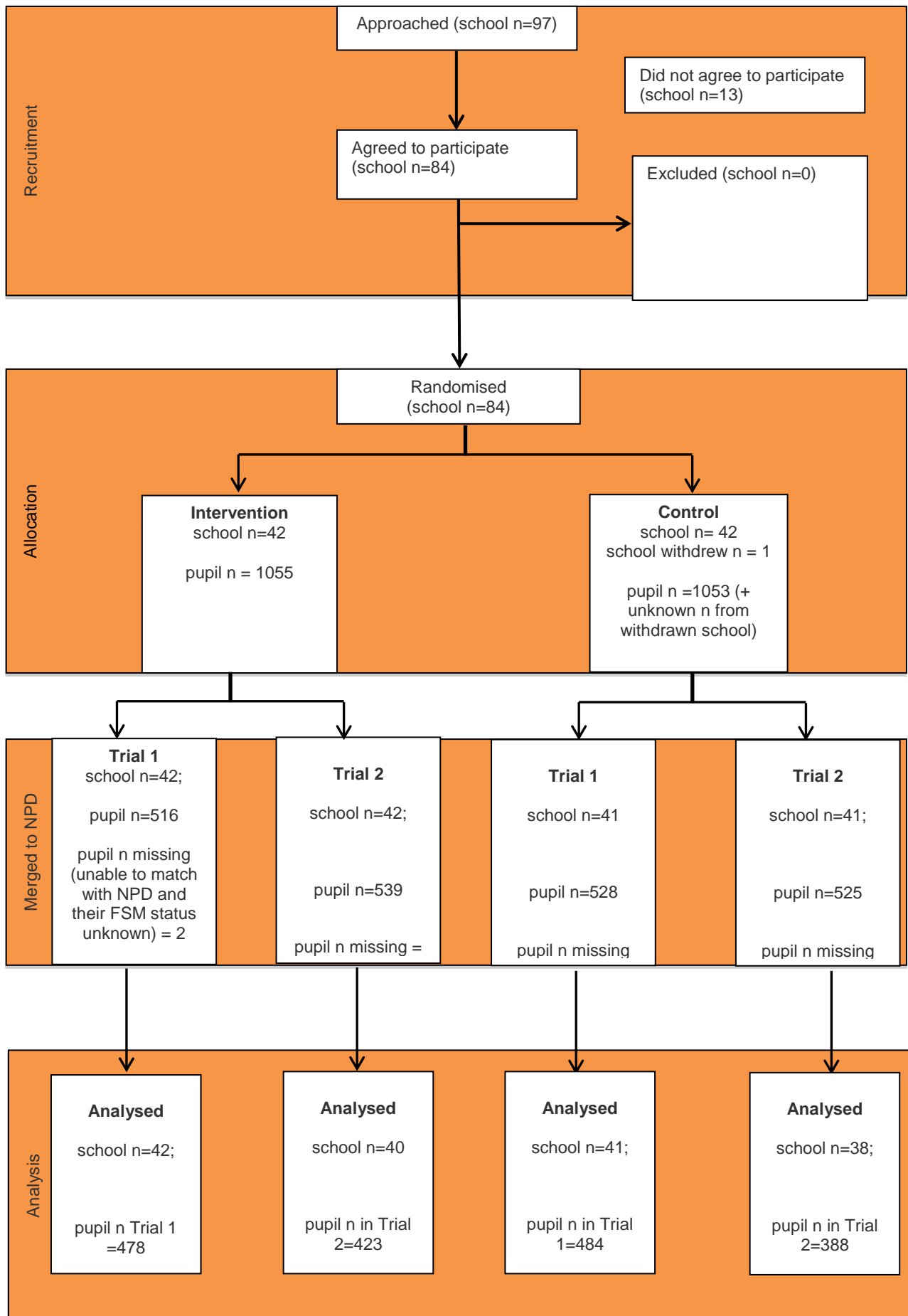
Seventy-eight schools agreed to facilitate the additional assessment to be used as the revised primary outcome.

Figures 1 and 2 represent participant flow in both trials for all pupils and FSM pupils respectively.

**Figure 1: Participant flow diagram—all pupils Trial 1 and Trial 2**



**Figure 2: Participant flow diagram—FSM pupils Trial 1 and Trial 2**





**Table 4: Minimum detectable effect size at different stages of Trial 1**

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
<b>MDES</b>		0.14	0.23	0.13	0.22	0.20	0.25
<b>Pre-test/ post-test correlations</b>	level 1 (pupil)	0.70	0.70	0.70	0.70	0.60	0.60
<b>Intracluster correlations (ICCs)</b>	level 2 (school)	0.08	0.08	0.08	0.08	0.20	0.25
<b>Alpha</b>		0.05	0.05	0.05	0.05	0.05	0.05
<b>Power</b>		0.8	0.8	0.8	0.8	0.8	0.8
<b>One-sided or two-sided?</b>		2	2	2	2	2	2
<b>Average cluster size</b>		45	9	45	11	31	12
<b>Number of schools</b>	intervention	40	40	42	42	42	42
	control	40	40	42	42	41	41
	<b>total</b>	80	80	84	84	83	83
<b>Number of pupils</b>	intervention	1800	375	1890	378	1243	478
	control	1800	375	1890	378	1222	484
	<b>total</b>	3600	750	3780	756	2465	962

**Table 5: Minimum detectable effect size at different stages of Trial 2**

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
<b>MDES</b>		0.15	0.22	0.13	0.20	0.20	0.26
<b>Pre-test/ post-test correlations</b>	level 1 (pupil)	0.70	0.70	0.70	0.70	0.63	0.60
<b>Intracluster correlations (ICCs)</b>	level 2 (school)	0.08	0.08	0.08	0.08	0.19	0.19
<b>Alpha</b>		0.05	0.05	0.05	0.05	0.05	0.05
<b>Power</b>		0.8	0.8	0.8	0.8	0.8	0.8
<b>One-sided or two-sided?</b>		2	2	2	2	2	2
<b>Average cluster size</b>		45	11	45	11	28	10
<b>Number of schools</b>	intervention	35	35	42	42	40	40
	control	35	35	42	42	38	38
	<b>total</b>	70	70	84	84	78	78
<b>Number of pupils</b>	intervention	1575	375	1890	462	1164	484
	control	1575	375	1890	462	1032	388
	<b>total</b>	3150	750	3780	924	2196	811

The minimum detectable effect sizes (MDES) at different stages of Trial 1 and Trial 2 are presented in Table 4 and Table 5 respectively. The MDES in each trial at analysis stage was bigger than the projected MDES at the protocol stage due to higher intra-cluster correlation, smaller pre-test and post correlation, and smaller total number of pupils. This could partly explain why the effect sizes for the primary outcomes in both trials were not statistically significant, due to lack of power. The estimated effect sizes are smaller than their expected MDESs at analysis stage. MDESs were calculated based on a cluster randomised trial with person-level outcomes using Optimal Design Plus Empirical Evidence software.

### School characteristics

Table 6 shows summary statistics of school-level characteristics of the 84 primary schools involved in the trial. The percentages of schools with mixed group classes were similar in the intervention and the control groups. Out of the 83 recruited schools, 44% (37) were from Leeds and 56% (47) were from Lincolnshire. The mean school size for the intervention group was 252 pupils with a range of 48 to 639 pupils. For the control group, it was 254 pupils with a range of 76 to 592. The median percentage of children eligible for FSM in the intervention group was 29.9% and 30.1% for the control schools.

**Table 6: School level characteristics**

Variable	Intervention group		Control group		
	School-level (categorical)	n/N (missing)	Percentage	n/N (missing)	Percentage
<b>Mixed age group classes</b>					
<b>Yes</b>	11/42 (0)	26.2%	10/41 (1)	24.4%	
<b>No</b>	31/42 (0)	73.8%	31/41 (1)	75.6%	
<b>Authority</b>					
<b>Leads</b>	18/42 (0)	42.9%	18/41 (1)	43.9%	
<b>Lincolns hire</b>	24/42 (0)	57.1%	23/41 (1)	56.1%	
School-level (continuous)	n (missing)	Median (range)	n (missing)	Median (range)	
<b>School size</b>	42 (0)	226 (48–639)	41 (1)	215 (76–592)	
<b>Percentage FSM</b>	42 (0)	29.9 (5.5–73.2)	41 (1)	34.1 (7.5–82.1)	

### Pupil characteristics

Baseline characteristics for all pupils involved in Trial 1 in Year 6 are presented in Table 7. There were 2,700 pupils from 83 primary schools (data not available for one school that withdrew before being informed of its random allocation) with approximately 50% of the pupils in the intervention group and control group, respectively. About 20% of the pupils in both the intervention and control groups had been eligible to receive free school meals at some point. The percentages of male and female pupils are similar in both groups. The average KS1 maths scores were  $15.5 \pm 3.7$  and  $15.4 \pm 3.8$  for the intervention and the control groups, respectively. The average KS1 reading scores were also similar in both groups at  $15.5 \pm 4.2$  and  $15.2 \pm 4.2$ , respectively. The average KS1 writing scores were  $14.3 \pm 3.9$  and  $14.3 \pm 4.0$ , respectively. This means that the academic performance of Year 6 pupils from the recruited schools was comparable at baseline based on their Key Stage 1 scores, although the randomisation was done at school level. The effect size for the pre-tests in trial was 0.03 (-0.18, 0.23), 0.05 (-0.16, 0.25), and -0.01 (-0.21, 0.20) for maths, reading, and writing, respectively. The distributions of the pre-test scores are presented in Figures 6–8 in Appendix R. The Key Stage writing scores appeared symmetric with similar number of pupils scoring below and above 15. The distribution of KS1 reading scores appeared skewed with more pupils scoring more than 15 than those that scoring less

than 15. The distribution of KS1 maths scores also appeared skewed. There is no evidence of ceiling or effect on the pre-test scores in Trial 1.

Baseline characteristics of all pupils involved in Trial 2 are also presented in Table 7. There were 2,817 pupils from 83 primary schools with approximately 52% of the pupils in the intervention group and 48% in the control group. Of these, 38% and 39.1%, respectively, had been eligible to receive free school meals at some point. The percentage on FSM in Trial 2 (~38.5%) was higher than those in Trial 1 (~20%). There were 53.2% males in the intervention group and 51.3% males in the control group. The average KS1 maths scores were  $15.7 \pm 3.4$  and  $15.4 \pm 3.7$  for the intervention and the control groups, respectively. The KS1 reading scores were also similar in both groups with averages of  $15.9 \pm 3.9$  and  $15.5 \pm 4.0$ , respectively. The average KS1 writing scores in the intervention group were  $15.5 \pm 3.7$  and  $14.4 \pm 4.0$ , respectively. Similar to Trial 1, the academic performance of the pupils was comparable at baseline based on their KS1 scores. The effect sizes for the pre-tests in trial were 0.07 (-0.09, 0.24), 0.1 (-0.16, 0.06), and 0.04 (-0.14, 0.21) for maths, reading and writing, respectively. The distributions of the pre-test scores are presented in Figures 4–5 in Appendix R. The distributions of the pre-test scores in Trial 2 were similar to the distributions of the pre-test scores in Trial 1, which is expected since the pre-tests in both trials are Key Stage 1 scores.

**Table 7: Pupil-level characteristics in Trial 1 and Trial 2**

	Variable	Intervention group		Control group		
		Pupil-level (categorical)	n/N (missing)	Percentage	n/N (missing)	Percentage
<b>Trial 1</b>	Eligible for FSM		269/1347 (2)	20.0	265/1335 (16)	19.9
	Gender					
	Female		642/1347 (2)	47.7	679/1335 (16)	50.9
	Male		705/1347 (2)	52.3	656/1335 (16)	49.1
<b>Trial 2</b>	Eligible for FSM		539/1418 (35)	38.0	525/1344 (20)	39.1
	Gender					
	Male		754/1418 (35)	53.2	689/1344 (20)	51.3

	Female	664/1418 (35)	46.8	655/1344 (20)	48.7	
	Pupil-level (continuous)	n (missing)	[Mean ± SD]	n (missing)	[Mean ± SD]	Effect Size
<b>Trial 1</b>	KS1 maths	1255 (84)	15.5 ± 3.7	1245 (106)	15.4 ± 3.8	0.03 (-0.18, 0.23)
	KS1 reading	1258 (91)	15.5 ± 4.2	1246 (105)	15.2 ± 4.2	0.05 (-0.16, 0.25)
	KS1 writing	1258 (91)	14.3 ± 3.9	1246 (105)	14.3 ± 4.0	-0.01 (-0.21, 0.20)
<b>Trial 2</b>	KS1 maths	1359 (94)	15.7 ± 3.4	1284 (80)	15.4 ± 3.7	0.07 (-0.09, 0.24)
	KS1 reading	1359 (94)	15.9 ± 3.9	1284 (80)	15.5 ± 4.0	0.1 (-0.06, 0.27)
	KS1 writing	1359 (94)	14.5 ± 3.7	1284 (80)	14.4 ± 3.7	0.04 (-0.14, 0.21)

Table 30 in Appendix S shows summary statistics for the final analysis sample. The two tables are very similar.

## Outcomes and analysis

### Primary analysis—impact of IPEELL on writing

The primary analyses were done under the intention-to-treat principle, which means all pupils were analysed according to the cluster randomisation of the schools. The primary outcome in Trial 1 was Key Stage 2 writing scores—ordinal data with seven levels. From highest to lowest, these are:

- GDS—‘working at greater depth within the expected standard’;
- EXS—‘working at the expected standard’;
- WTS—‘working towards the expected standard’;
- PGK—‘growing development of the expected standard’;
- PKE—‘early development of the expected standard’;
- PKF—‘foundation for the expected standard’; and

- BLW—‘below the standard of the pre-key stage’.

As shown in Table 8, 70% and 73% of the pupils in the intervention and control groups were rated as ‘working at the expected standard’ or at greater depth. Analysis of the ordinal data using a robust proportional odds model shows that the odds ratio of having higher writing scores was 16% less in the intervention group than in the control group. The implication of this is that pupils in the intervention group were not doing better than those in the control group. In order for the result to be comparable to effect size, the proportional odds ratio was converted to effect size as suggested by Borenstein *et al.* (2009). Note: due to the large sample size, Cohen’s *d* and Hedges’ *g* effect sizes are equivalent because the multiplicative correction factor in Hedges’ *g* effect size is approximately 1. The equivalent effect size of -0.09 (-0.34, 0.15) shows that Year 6 pupils from the intervention schools were not doing better in writing than Year 6 pupils from the control group.

**Table 8: Analysis of the primary outcome for Trial 1 and Trial 2**

The ordinal writing scores in Trial 1 were analysed using a robust proportional odds model. Trial 2 was analysed using a multilevel model.

	Raw means				Effect size			
	Outcome	Intervention group		Control group		n in model (intervention; control)	Proportional Odds Ratio	Hedges g (95% CI)
		n (missing)	Proportion (95% CI)	n (missing)	Proportion (95% CI)			
<b>Trial 1</b>	Writing (Expected Standard or higher)	1243 (106)	0.70 (0.68, 0.73)	1222 (129)	0.73 (0.71, 0.76)	2465 (1243, 1222)	0.84 (0.54, 1.32)	-0.09 (-0.34, 0.15)
	Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	n in model (intervention; control)	Hedges g (95% CI)	
<b>Trial 2</b>	Writing	1164 (289)	21.6 (21.4, 22.0)	1032 (332)	20.9 (20.5, 21.2)	2196 (1164, 1032)	0.11 (-0.13, 0.34)	

The writing outcomes for Trial 1 (Table 9) were treated as scores to compute standard deviation and mean difference as required by the EEF reporting template. The unadjusted mean difference in Trial 1 was -0.05 (-0.16, 0.06), whilst the mean difference adjusted for the minimisation factors was -0.05 (-0.17, 0.07). In Trial 2, the unadjusted mean difference was 0.64 (-0.65, 1.93), whilst the mean difference after adjusting for the minimisations factors was 0.60 (-0.56, 1.75).

Table 9 shows the mean differences between the IPEELL intervention and the comparison group. The adjusted differences in means denote the mean difference after adjusting for the minimisation factors (school size, local authority, mixed classes, and FSM). Unconditional effect sizes, using unconditional pooled standard deviation, are -0.07 (-0.24, 0.10) in Trial 1 and 0.08 (-0.08, 0.25) in Trial 2. These effect sizes are slightly different from the effect sizes from the multilevel model approach of Hedges (2007). For example, the unconditional effect size in Trial 2 is 0.08 (-0.08, 0.25) which is slightly smaller and has narrower 95% confidence intervals than our reported effect size of 0.11 (-0.13, 0.34).

However, unconditional effect size of -0.07 (-0.24, 0.1) in Trial 1 is slightly bigger and has narrower 95% confidence intervals than the multilevel effect size of -0.09 (-0.33, 0.15). The proposed unconditional effect size of dividing adjusted mean difference by unconditional pooled variance is ad-hoc in nature and dividing 95% confidence intervals of mean difference by a constant unconditional pooled variance is prone to narrower confidence intervals because it assumes the unconditional variance is constant. The correct approach is treating effect size as a random variable with its own distribution, which is the approach proposed by Hedges (2007) and reported in our report.

**Table 9: Estimation of unconditional effect sizes in Trial 1 and Trial 2**

Effect size was calculated by dividing adjusted mean difference and its 95% confidence intervals by the unconditional pooled standard deviation

	Outcome	Unadjusted differences in means	Adjusted differences in means	IPEELL group		Control group		Pooled variance	Unconditional effect size
				n (missing)	Variance of outcome	n (missing)	Variance of outcome		
<b>Trial 1</b>	Writing (ordinal scores)	-0.05 (-0.16, 0.06)	-0.05 (-0.17, 0.07)	1243 (106)	0.5	1222 (129)	0.5	0.5	-0.07 (-0.24, 0.10)
<b>Trial 2</b>	Writing	0.64 (-0.65, 1.93)	0.60 (-0.56, 1.75)	1164 (289)	47.5	1032 (332)	52.6	50.5	0.08 (-0.08, 0.25)

To further understand the results of Trial 1, Table 1517 in Appendix O shows the proportion of pupils according to the levels of the ordinal writing grades. In the intervention group, 9.1% of pupils were rated to be working at greater depth, compared with 10.4% in the control group. In the intervention and control groups, 59.4% and 60.7, respectively, were rated to be working at the expected standard. However, 31.5% and 28.9% of the pupils in the intervention and the control groups were rated to be below the expected standard. The proportional odds model assumed a constant effect of the intervention for adjacent levels of the writing outcome, which is a drawback of the model since intervention may have different effect in pupil working below the expected standard and those at expected standard or higher. The ordinal writing outcome was also analysed using a multilevel model by assigning scores to the different levels. The ordinal levels were assigned scores from one to seven with score of seven for 'working at greater depth within the expected standard', score of six for 'working at expected standard', until a score of 1 for 'working below the standard of the pre-key stage'. The estimated effect size from the multilevel model presented in

Table 24 in Appendix P was -0.09 (-0.33, 0.15), which is similar to the effect size presented in Table 8, which shows that the results are most likely due to the data rather than the choice of analytical methods.

Unlike in Trial 1, the primary analysis in Trial 2 was a trial-specific writing test, which was a test that drew on components of historical KS2 assessments (past KS2 writing papers). Seventy-eight out of the 83 schools took part in the test, which means that about 17.4% of 2,762 pupils in Trial 2 merged to National Pupil Database did not provide outcome data. The average writing score in the intervention

group was 21.6, and 20.9 in the control group. There was no evidence of a positive impact of the two-year IPEELL intervention on the writing outcome with an effect size of 0.11 (-0.13, 0.34).

Table 21 in Appendix O presents the regression coefficients and the variance components from the multilevel analysis of the primary outcome in Trial 2. The mean difference between the intervention and the control groups was 0.60 (-0.56, 1.75). The intra-cluster correlation was 0.19. To check the impact of mixed classes in Trial 2, the data was analysed without schools with mixed Year 5 and Year 6 classes. The effect size was 0.14 (-0.43, 0.41), which is not substantially different from the effect size obtained from the analysis of all the 78 schools that took part in the study-specific writing test. Further analysis checked the impact of the minimisation factors by only adjusting for Key Stage 1 writing scores. The result presented in Table 16 in Appendix O shows an effect size of 0.11 (-0.13, 0.34), which is the same as the effect size presented in Table 8. This means that in this specific case, adjusting for minimisation factors makes little difference from only adjusting for Key Stage 1 writing scores. The effect size of -0.09 (-0.35, 0.16) from adjusting for only for Key Stage 1 writing scores is also similar to the effect size reported for Trial 1 in Table 8. This means that minimisation factors do not influence the effect sizes reported for both trials, which is expected since the factors are comparable at baseline.

### **Subgroup analysis of primary outcome—impact of IPEELL on writing**

Subgroup analyses of the primary outcomes by FSM, low achievers, and high achievers are presented in Table 10. The percentages of pupils on free school meals in Trial 1 that were rated to be working at the expected standard or at greater depth were 60% and 63% for the intervention and the control groups, respectively. The results also show a negative impact of IPEELL on this subgroup of pupils. The percentages of pupils classified as low achievers that were rated to be working at the expected standard or greater depth were 40% and 45% for the intervention and the control groups, respectively. The percentages of pupils classified as high achievers that were rated to be working at the expected standard or greater depth were 90% for both the intervention and the control groups. Table 8 shows a negative impact of IPEELL on writing outcomes. The proportions of pupils according to the levels of the writing outcome are presented in Tables 17–20, Appendix O.



**Table 10: Subgroup analysis of the primary outcome in Trial 1 and Trial 2 by FSM, low achievers, and high achievers**

Trials	Raw means					Effect size		
	Outcome	Intervention group		Control group		n in model (intervention; control)	Proportional Odds Ratio	Hedges g (95% CI)
		n (missing)	Proportion (95% CI)	n (missing)	Proportion (95% CI)			
Trial 1	FSM	478 (38)	0.60 (0.56, 0.64)	484 (44)	0.63 (0.59, 0.68)	962 (478, 484)	0.87 (0.50, 1.51)	-0.07 (-0.38, 0.23)
	Low achievers	494 (102)	0.40 (0.36, 0.45)	454 (117)	0.45 (0.40, 0.49)	948 (494, 454)	0.79 (0.47, 1.35)	-0.13 (-0.42, 0.16)
	Higher achievers	749 (4)	0.90 (0.88, 0.92)	768 (12)	0.90 (0.88, 0.92)	1517 (749, 768)	0.96 (0.59, 1.56)	-0.02 (-0.29, 0.25)
Trial 2	FSM	423 (116)	20.2 (19.5, 20.9)	388 (137)	18.6 (17.9, 19.3)	811 (423, 388)	0.14 (-0.14, 0.41)	
	Low achievers	435 (101)	17.38 (16.8, 17.9)	327 (153)	15.1 (14.5, 15.7)	762 (435, 327)	0.26 (-0.01, 0.53)	
	Higher achievers	729 (94)	24.1 (23.7, 24.6)	705 (99)	23.6 (23.1, 24.1)	1434 (729, 705)	0.06 (-0.21, 0.33)	

In Trial 2, there were 811 FSM pupils, with 52% in the intervention group and 48% in the control group. The mean writing scores were 20.2 in the intervention group and 18.6 in the control group. The two-year IPEELL intervention had a positive impact on the writing outcome for this subgroup of pupils with an effect size of 0.14 (-0.14, 0.41), which was not statistically significant. The two-year IPEELL intervention also had a bigger impact on low achieving pupils compared to high achieving pupils. The effect size for low achievers was 0.26 (-0.01, 0.53), which was marginally non-significant, whilst the effect size for high achievers was 0.06 (-0.21, 0.33).

**Secondary analyses—impact of IPEELL on maths, reading, general punctuation and spelling (GPS), and writing**

The secondary outcomes in Trial 1 were analysed using multilevel models to account for intra-school correlation. The results are presented in Table 11. The average post-intervention maths scores were 101.69 and 102.27 for the intervention and the control groups, respectively. The effect size of -0.22 (-0.47, 0.04) implies that IPEELL did not improve the maths outcomes of Year 6 pupils in Trial 1. The average post-intervention reading score was 100.85 in the intervention group, and 101.91 in the control group. The effect size was -0.23 (-0.43, -0.04). The significant effect size implies that IPEELL had a negative impact on reading outcomes of Year 6 pupils in Trial 1. Finally, the average post-intervention general punctuation and spelling was 102.71 in the intervention group, and 103.44 in the control group. The effect size was -0.22 (-0.46, 0.01), which also indicates that IPEELL did not improve the general punctuation and spelling outcomes and is close to a significant negative difference.

Analyses of the secondary outcomes in Trial 2 are also presented in Table 11. The average post-intervention maths scores were 102.5 in the intervention group, and 103.1 in the control group. The effect size was -0.30 (-0.53, -0.02), which shows that two year IPEELL intervention did not improve the Key Stage 2 maths results of these pupils. Similarly, the intervention had an exaggerated negative impact on Key Stage 2 reading and the general punctuation and spelling results. The effect size for reading was -0.17 (-0.33, -0.01) and -0.28 (-0.49, -0.06) for general punctuation and spelling. The Key Stage 2 writing scores (national test), which was an ordinal dataset, was analysed using robust proportional odds model; 70.2% of the pupils in the intervention group and 71.2% of those in the control group were rated to be 'working at greater depth within the expected standard' or 'working at the expected standard'. The proportional odds ratio was 0.48 (0.39, 0.58) with a corresponding effect size of -0.04 (-0.25, 0.17). The intervention consistently shows negative impacts on the outcomes.

**Table 1111: Analysis of secondary outcomes in Trial 1 and Trial 2**

Trial	Raw means					Effect size	
	Outcome	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)
		n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)		
Trial 1	Maths	1226 (123)	101.69 (101.29, 102.09)	1215 (100)	102.67 (102.27, 103.08)	2441 (1226, 1215)	-0.22 (-0.47, 0.04)
	Reading	1220 (129)	100.85 (100.38, 101.32)	1209 (142)	101.91 (101.43, 102.38)	2429 (1220, 1209)	-0.23 (-0.43, -0.04)
	Spelling	1228 (121)	102.71 (102.28, 103.13)	1222 (129)	103.44 (103.01, 103.87)	2450 (1228, 1222)	-0.22 (-0.46, 0.01)
Trial 2	Maths	1326 (127)	102.5 (101.8, 102.6)	1226 (138)	103.1 (102.6, 103.5)	2552 (1326, 1226)	-0.30 (-0.53, -0.07)
	Reading	1328 (125)	102.2 (101.8, 102.7)	1221 (143)	102.7 (102.2, 103.2)	2549 (1328, 1221)	-0.17 (-0.33, -0.01)
	Spelling	1328 (125)	104.2 (103.7, 104.6)	1228 (136)	105.0 (104.5, 105.4)	2566 (1328, 1228)	-0.28 (-0.49, -0.06)
	Writing (Expected Standard or higher)	1411 (42)	70.2 (67.8, 72.6)	1328 (36)	71.2 (68.7, 73.6)	2739 (1411, 1328)	0.48 (0.39, 0.58)

### Subgroup analysis of secondary outcomes—impact of IPEELL on maths, reading, general punctuation and spelling (GPS), and writing

Subgroup analyses of secondary outcomes in Trial 1 are presented in Table 12. The average post-intervention maths scores for pupils on free school meals were 99.61 and 100.24 for the intervention and the control groups, respectively. The effect size of -0.16 (-0.44, 0.12) implies that IPEELL did not improve maths outcome of Year 6 FSM pupils. The average post-intervention reading scores for FSM pupils were 98.24 and 99.46 for the intervention and the control groups, respectively. The effect size of -0.22 (-0.46, 0.02) implies that IPEELL did not improve reading outcomes of FSM pupils. The average

post-intervention general punctuation and spelling scores for FSM pupils were 100.38 and 100.75 for the intervention and the control groups, respectively. The effect size of -0.13 (-0.40, 0.13) implies that IPEELL did not improve the general punctuation and spelling. Subgroup analyses of secondary outcomes by low and high achievers are presented in Tables 23–24 in Appendix P. The effect sizes of low achievers were -0.24 (-0.53, 0.05), -0.28 (-0.55, -0.02), and -0.24 (-0.52, 0.05) for maths, reading, and spelling respectively. For high achievers, the effect sizes were -0.18 (-0.45, 0.08), -0.19 (-0.38, -0.01), and -0.18 (-0.42, 0.06), respectively for maths, reading, and spelling. The results show consistently that IPEELL did not improve the performance of the pupils in Trial 1 and (at least for reading) had a negative impact.

**Table 12: Analysis of the secondary outcomes in Trial 1 and Trial 2—FSM pupils**

Trial	Outcome	Raw means				Effect size		
		Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	
		n (missing)	Mean (95% CI)	N (missing)	Mean (95% CI)			
Trial 1	Maths	469 (47)	99.61 (98.95, 100.27)	481 (47)	100.24 (99.59, 100.89)	950 (469, 481)	-0.16 (-0.44, 0.12)	
	Reading	464 (52)	98.24 (97.46, 99.03)	472 (56)	99.46 (98.68, 100.24)	936 (464, 472)	-0.22 (-0.46, 0.02)	
	Spelling (GPS)	470 (46)	100.38 (99.67, 101.09)	483 (45)	100.75 (100.05, 101.45)	953 (470, 483)	-0.13 (-0.40, 0.13)	
Trial 2	Maths	498 (41)	100.4 (99.7, 101.1)	473 (52)	100.6 (99.8, 101.3)	971 (498, 473)	-0.25 (-0.52, 0.02)	
	Reading	498 (41)	100.3 (99.5, 101.1)	470 (55)	100.1 (99.3, 100.9)	960 (498, 470)	-0.13 (-0.33, 0.07)	
	Spelling	498 (41)	102.6 (101.8, 103.3)	475 (50)	102.3 (101.5, 103.0)	973 (498, 475)	-0.18 (-0.43, 0.07)	
	Writing (Expected Standard or higher)	532 (9)	62.4 (58.2, 66.4)	516 (9)	59.1 (54.8, 63.3)	1046 (532, 516)	1.00 (0.64, 1.57)	0.0 (-0.25, 0.25)

Subgroup analyses of the secondary outcomes in Trial 2 are also presented in Table 12. The average post-intervention maths scores for FSM pupils was 100.4 in the intervention group, and 100.6 in the control group. The effect size of -0.25 (-0.52, 0.02) indicates that the two years of IPEELL intervention did not improve maths outcome of FSM pupils. The average post-intervention reading scores for FSM pupils were 100.3 and 100.1 for the intervention and the control groups, respectively. The effect size of -0.13 (-0.33, 0.07) also means that the two-year IPEELL intervention did not improve their reading outcomes. Similarly, the effect size of -0.18 (-0.43, 0.7) for general punctuation and spelling, and 0.0 (-0.25, 0.25) for Key Stage 2 writing (national test) imply non-positive impact of the intervention. Subgroup

analyses by low and high achievers are presented in Table 2023 and Table 2124 in Appendix P. The effect sizes for the low achiever group were -0.29 (-0.56, -0.01), -0.11 (-0.30, 0.08), -0.23 (-0.49, 0.03) and 0.01 (-0.25, 0.27) for KS2 maths, reading, general punctuation and spelling, and writing (national test), respectively. For the high achiever group, the corresponding effect sizes were -0.29 (-0.57, -0.02), -0.17 (-0.36, 0.01), -0.27 (-0.51, -0.03) and -0.05 (0.27, 0.18). The results show a consistent pattern of no significant positive impact of the two-year IPEELL intervention on pupils' KS2 results and suggest a negative impact.

### **Missing data analysis for primary outcome—impact of IPEELL on writing**

The percentage of missing data at pupil level is 9.0% in Trial 1—8.5% of the pupils in the intervention group and 9.5% in the control group. Although it was stated in the protocol that multiple imputation will be performed to sensitise missing data, the dropout model presented in Table 25 in Appendix Q shows no association between missingness and the baseline factors.

There was higher proportion of pupils with missing data in Trial 2. This is due to the fact that only 78 out of the 83 schools took part in the study-specific writing test. This means that 20.1% of 2,762 pupils in Trial 2 merged to National Pupil Database did not provide outcome data.

The drop out model presented in Table 26 in Appendix Q shows a significant association between missing outcome and Key Stage 1 writing score, Local Authority (Leeds or Lincolnshire) and FSM status. It is important to note that it is tricky to attribute the missing data to pupils' characteristics since it was the whole school that did not take part in the study-specific writing test for Trial 2. As a sensitivity analysis, the missing outcomes in the writing outcomes were imputed using multiple imputation technique for missing data. The results, presented in Table 27 in Appendix Q, show that the effect size with and without multiple imputations are comparable. The effect size from multiple imputations of all pupils was 0.11 (-0.09, 0.27). The effect sizes for the pre-specified subgroup analyses were 0.12 (-0.11, 0.32), 0.20 (-0.01, 0.39) and 0.05 (-0.17, 0.25) for FSM, low achievers, and high achievers, respectively.

### **CACE Analysis—impact of IPEELL on writing**

The quality of implementation, assessed in spring and autumn 2016, was used as a measure of compliance for Trial 1 and Trial 2, respectively. In spring 2016 (the assessment of Trial 1):

- 14.3% (6/42) of the intervention schools were rated as 'very good' for the implementations of the intervention;
- 45.2% (19/41) were rated as 'good';
- 35.3% (14/41) as 'satisfactory'; and
- 4.9% (2/41) as 'unsatisfactory'.

In autumn 2016 (the assessment of Trial 2):

- 19.5% (8/41) of the intervention schools were rated as 'very good' for the implementations of the intervention;
- 39.0% (16/41) were rated as 'good';
- 28.3% (12/41) as 'satisfactory'; and
- 12.2% (5/41) as 'unsatisfactory'.

One of the schools in the intervention group had no quality of implementation score. We considered two scenarios to estimate the impact of quality of implementation on the resulting effect sizes for the primary outcome under the intention-to-treat principles. In the first scenario, compliance was defined as 'very good' or 'good' implementation and non-compliance was defined as 'satisfactory', 'unsatisfactory', or no intervention. In the second scenario, compliance was defined as 'very good', 'good' or 'satisfactory' implementation, whilst non-compliance was defined as 'unsatisfactory' or no intervention.

We assumed no school in the comparison group received the IPEELL intervention. The CACE analysis used the approach of Schochet and Chiang (2011) where CACE effect size (CACE ES) was estimated by weighting the effect size under intention to treat by the proportion of compliance. The CACE results, presented in Table 13, show that the impact of the IPEELL intervention on writing outcomes in Trial 2 could be up to 0.19 (-0.23, 0.60) had the intervention been implemented up to good level. The CACE effect size for Trial 1 is counterintuitive because it implies good implementation of the intervention worsens the impact of the intervention. Table 28 in Appendix S presents exploratory analysis of the impact of the quality of implementation of IPEELL interventions by comparing the four categories of quality of implementation with the comparison group. For Trial 2, schools with very good implementation of the intervention had an effect size of 0.30 (-0.08, 0.69) compared to an effect size of 0.05 for schools with unsatisfactory implementation of the intervention.

**Table 13: Exploratory CACE analysis to investigate the impact of the quality of IPEELL implementation on the estimated effect sizes**

Trials	'Good' implementation			'Satisfactory' implementation		
	% compliance	ITT ES	CACE ES	% compliance	ITT ES	CACE ES
1	59.5%	-0.09 (-0.34, 0.15)	-0.15 (-0.57, 0.25)	92.9%	-0.09 (-0.34, 0.15)	-0.10 (-0.37, 0.16)
2	57.1%	0.11 (-0.13, 0.34)	0.19 (-0.23, 0.60)	88.1%	0.11 (-0.13, 0.34)	0.12 (-0.15, 0.39)

### Deviation from Statistical Analysis Plan

There are two major deviations from the Statistical Analysis Plan. The primary analysis stated in the SAP is a model with only pre-test and intervention as the covariates, but the primary analyses reported adjusted for minimisation factors in accordance with the updated EEF Statistical Analysis Guidance. The second deviation from SAP is that quality of implementation was considered as proxy for compliance in the Complier Average Causal Effect (CACE) analysis. It was not possible to collect compliance data on pupil's participation because of the nature of the intervention.

### Cost

Detail on costs was provided by CEP. The costs to schools of repeating the programme is estimated at:

- one day's training—venue, equipment, refreshments, and training materials: £900;
- IPEELL trainers x2, with one day's preparation, £600 per trainer: total £1,800; and
- 'memorable experiences' per class: £990.

The following assumptions are built into the calculations for implementation by other schools:

- IPEELL training delivered to 30 teachers; and
- IPEELL delivered to a class of 30 pupils.

Basic costs to schools would then be £90 per teacher for the initial training and £33 per pupil for memorable experiences, that is, £990 per class. These costs could be reduced significantly depending

on the cost of the memorable experiences chosen; they would increase if supply cover for the teacher training was required.

Additional resources required would be minimal as the strategies can be delivered as an integral part of literacy/English teaching. The project team has developed a toolkit. A school would need to buy a minimum of one toolkit, but preferably one per teacher implementing IPEELL and one for the literacy co-ordinator. Costs in future years would relate to the cost of the chosen memorable experiences and any further training required for teachers new to the school.

The cost of the approach is thus estimated at £36 per pupil. This estimate includes training and materials (£90 per teacher or £3 per pupil), and the cost of memorable experiences (£33 per pupil). Estimates are based on a class of 30 pupils, and on training being delivered to a group of 30 teachers.

**Table 14: Projected costs per pupil over a three year period**

Group	Year 1	Year 2	Year 3
Training costs of teachers (including materials)	£90	£0	£0
Memorable experiences	£990	£990	£990
Cost per pupil per year over three years	£36	£33	£33

Table 14 assumes the same class teacher delivers the intervention over a three year period and requires no top-up training. This means that for years two and three schools only require the cost to cover the memorable experiences.

## Implementation and process evaluation

### Implementation

#### What are the necessary conditions for success of the intervention?

This section presents and discusses the teacher manual, the training and networking meetings, and features of the delivery of the intervention as the necessary conditions for the success of the intervention.

#### Teacher manual

In our judgement, the teacher manual is of exceptional quality. It is clear; it contains numerous illustrative examples including engaging tables, graphs, and photos and guides the reader through a clear progression of ideas relevant to the different stages of the writing process.

All teachers in the schools visited agreed that the teacher manual was particularly valuable as a document to constantly refer back to and helped to fill any gaps in understanding the IPEELL writing process after the training. It was thought to be an 'accessible' document that could provide essential guidance at a 'glance' (School 58, Summer 2015). It also allowed teachers to share ideas and aspects of IPEELL with other teachers or teaching assistants within the school that did not have the opportunity to attend the training (for example, School 50, Spring 2016). The CD-ROM, in particular, was very useful to many teachers as the available resources provided them with a starting point of material which could then be adapted to work within each school context and according to the focus of the lesson and the children's needs.

Many teachers mentioned that they would have liked to have been provided with a greater range of model texts. Firstly, these were thought to be useful as example texts for the Year 5 children moving up to Year 6 (for example, School 59, Spring 2016) as it would save teachers time in trying to develop their own materials. In a control school (School 52) visited in spring 2017 (the control group schools were waitlisted and received the initiative in the second year of the trial), most writing examples were used with Years 3 and 4 and the school would have liked to have more available for Years 5 and 6:

*'We are having to generate more really, so we need 4 or 5. Because every time we do a piece of writing, we are modelling something different, similar to it, and every time it needs somebody to write a full piece, to plan it out, to write it. It would have been good to have (more model texts) and really stepping up the levels of what a really good one looks like rather than having to write our own'* (teacher, School 19).

As teachers were adapting IPEELL to their school context and their children's needs, more examples of adapted planning frames or marking sheets would have been beneficial (School 50, Spring 2016).

#### Training and networking meetings

All teachers in all schools commented favourably on the training. The aspects of the training that dealt with guidance, demonstration, and modelling were perceived as being particularly beneficial. Furthermore, the networking meetings with other IPEELL schools that followed the initial training were felt to be very useful as they enabled teachers to share ideas around implementation, and, in one case (School 50, Spring 2016), led to teachers reconsidering aspects of their practice. After a networking meeting, they started thinking more about the importance of motivational messages and tried to 'find manageable ways and meaningful ways for the children to be able to do those'. The meeting helped broaden their thinking about the nature of memorable experiences and sowed the seeds for a memorable Indian day experience:



*'Listening to some of other people's experiences gave us the idea to do this. We're gonna stop at the gate and not let them in the school and we're going to take them on a walk for the same length of time that a child in a village would need to get to school and then we're going to have a whole day setting it up as it would be in an Indian school. It's not going to cost a huge amount but they will remember it and they are going to get something different to write about.'*

## The delivery of the intervention

### Quality of the IPEELL delivery

Out of the 16 visits to intervention schools conducted from autumn 2015 to summer 2017, eight writing lessons using IPEELL were judged as being 'good' in terms of the aims and objectives of IPEELL, six were deemed to be 'satisfactory', and two were assessed as being 'unsatisfactory' (in one of these two schools, two observations were conducted, one of which was deemed 'unsatisfactory' and the other 'satisfactory'). In this section, we attempt to unpack and elaborate on the quality of the delivery of IPEELL elements that led to some lessons being deemed as 'good' or 'satisfactory' and others as 'unsatisfactory'. These elements form part of Table 15 (see below). They were chosen as a checklist of the quality of delivery because they directly match onto the aims and objectives of IPEELL which are in turn reflected onto the training materials. In other words, we aim to explore what aspects of the delivery led to this differential quality. Examples from particular observations are used to illustrate the points made.

As shown in Table 15 (below), reference to a **memorable experience** was common in most of the 'good' IPEELL writing observations but less so in those lessons deemed as satisfactory or unsatisfactory (7/8 for the 'good' lessons, 3/6 for the 'satisfactory' lessons and 1/2 for the 'unsatisfactory' ones). In this last school, we carried out two different lesson observations). For example, in one of the 'good' lesson observations, the children paid a visit to a Museum. During this visit, the children had a chance to handle and explore a variety of ancient Egyptian objects; they did a trail where they had to complete different challenges using their comprehension skills and they had different interactive workshops. The IPEELL writing lesson was based around these experiences.

Moreover, good or satisfactory **links between the writing activity and IPEELL** were seen in all 'good' lesson observations, in the majority of the 'satisfactory' ones (apart from one lesson) and in only one lesson out of the two visits that was deemed to be 'unsatisfactory'. In School 22, the writing activity was not linked explicitly to the IPEELL strategy. The lesson was part of the children's introductory work on the topic of a newspaper investigation. Pupils were given extracts of news reports. They were asked to read the articles provided, identify common features, and make observations about the language and tone used. They then cut out interesting extracts and presented these visually in the form of a poster. In School 16, pupils wrote down ideas/keywords/descriptive language that they would then use to write their myth. This was the first stage of their planning and they were then asked to use their best ideas to complete the IPEELL planning frame. Some went straight into developing their 'Points' (this term refers to developing key ideas which would later be further elaborated and exemplified in the writing) and filling in the relevant boxes. However, the initial box that referred to the mnemonic PAT (this required the pupils to decide what is the Purpose, Audience, and Type of presentation being used for the particular writing genre) was mostly left empty: when asked, pupils could not articulate or seem to understand the importance of writing for a particular purpose and a specific audience.

Furthermore, it was more likely for instances of **teacher modelling in the genre** to be seen in the 'good' lesson observations (8/8) compared to the 'satisfactory' (3/6) and the 'unsatisfactory' ones (0/2). For example, in School 19, good modelling was taking place on the board by the teacher and the Points part of the planning frame was being developed as directed by the teacher. In contrast, in School 53, the task that the children were working on was a 'cold' task as mentioned by the teacher. The pupils would then evaluate that cold task but it was not evident whether there would be any form of guided

support to then write a 'hot' task. Using the IPEELL marking criteria without using the planning frame or doing any teaching on IPEELL was confusing. Additionally, **teacher modelling using the planning frame** was more often observed in the 'good' lessons. For example, in Schools 81, 11A, 50A, and 59, the teacher deconstructed a pre-written text on the board to model good writing principles, while in School 19, the teacher modelled IPEELL writing developing the Points part of the planning frame on the board. However, it needs to be recognised that lack of evidence of teacher modelling in the lesson observations did not necessarily indicate a weakness of the implementation of IPEELL as it was often the case that teacher modelling took place in a previous lesson or during the first months of implementing the strategy in the schools.

**The planning frame was used appropriately by the children** in all 'good' and 'satisfactory' lessons. In some lesson observations (see section below, IPEELL becoming embedded), especially towards the end of the first year of the implementation of the strategy, it was observed that the planning frame had started fading out which was a sign that the strategy was becoming well embedded. This was mostly the case with some of the 'good' lessons (for example, School 11B, 50B, 81). In the two 'unsatisfactory' lessons, either the children did not use a planning frame (School 53) or its use was confusing (School 82). In School 53, for example, an IPEELL mark sheet was used to assess a cold task but there was no evidence of using the IPEELL planning frame. In School 82, the IPEELL planning frame was used at a second planning stage after pupils had already noted some preliminary planning ideas. This seemed to cause confusion in the children.

**Marking using the IPEELL scoring sheet** adapted for the different writing genres and children's ability was done appropriately and comfortably in all 'good' and 'satisfactory' lessons. In contrast, IPEELL marking was done inappropriately in the two 'unsatisfactory' lessons. In School 53, the pupils used an IPEELL mark sheet to evaluate their cold task but there was no evidence whether there would then be any form of guided support to write a hot task using the IPEELL planning frame. Similarly, in School 82, the assessment criteria used for self-scoring did not reflect the IPEELL whole piece structure. An assessment chart was distributed to the children where they were asked to provide a score for ten different criteria. However, it was not evident how these criteria would be considered in relation to the whole writing piece or for particular paragraphs. In relation to teacher scoring, there was some evidence of this in the children's books, mainly in the form of numbers. Nevertheless, there were no criteria included so it was not easy to understand what exactly these numbers indicated in terms of the children's writing performance.

As shown in the following table (Table 15), self-marking was used consistently and appropriately in all schools. Self-mark sheets were developed according to the different genres and the children's ability. The children seemed comfortable with the marking process as most teachers were careful to discuss the assessment criteria in the classroom making sure that these were well understood. Peer marking was used in most schools apart from School 16 where no evidence of peer marking was seen and Schools 19, 50A, 22, and 27 where peer marking was occasionally used but not consistently (indicated in the table by the symbol  $\sqrt{?}$ ). In School 19, for example, teachers felt that peer marking was taking up too much valuable time which could be devoted to 'getting another piece of writing done', as one teacher said. Furthermore, graphing of scores was evident in the children's books in most schools but, as discussed below (in the section Graphing of Scores—Pupils' Perceptions), the majority of schools did not continue with this as pupils gained confidence in writing and became more independent writers. Another important aspect of self-regulation, **goal setting**, was encouraged across all schools but it was less reinforced in four of the schools visited (Schools 19, 58, 27, and 82). In these schools, goal setting was recognised as an essential element of the strategy but there was limited evidence of it in the children's books. In School 82, for example, targets were mainly set by the teacher as was also the case in School 19. Moreover, in the latter case, pupils mentioned that they did not consider themselves able to evaluate their own work and come up with an accurate writing goal as they thought that the teacher was the expert in doing so.

Last but not least, there was very good evidence of the use of motivational messages across the schools visited. In most schools, pupils were encouraged to develop their personal motivational messages which were often displayed on classroom walls and in the children's books. In two schools (Schools 22 and 82), this was an area that was being developed at the time of our visit.

**Table 15: Checklist of the quality of delivery of IPEELL elements across the schools visited**

\* First visit to the school; \*\* second visit to the school.

Schools	Good						Satisfactory						Unsatisfactory			
	11B*	50B*	81	33	19	11A*	50A**	59	22	48	23	60	58	27	53	82
<i>Memorable experience</i>	√	√	√	√	x	√	√	√	x	x	√	x	√	√	√	x√
<i>Good links with the IPEELL strategy</i>	√	√	√	√	√	√	√	√	x	√	√	√	√	√	x	x√
<i>Teacher modelling in the genre</i>	√	√	√	√	√	√	√	√	x	x	√ poetry	√	√	x	x	xx
<i>Children using planning frame</i>	√	√	√	√	√	√	√	√	√	√	√	√	√	√	x	x
<i>Use of IPEELL scoring sheet</i>	√	√	√	√	√	√	√	√	√	√	√	√	√	√	x	x
<i>Self-scoring</i>	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
<i>Peer scoring</i>	√	√	√	√	√?	√	√?	√	√?	x	√	√	√	√?	√	x
<i>Goal setting</i>	√	√	√	√	x?	√	√	√	√	√	√	√	x?	x?	√	x?
<i>Motivational messages</i>	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√

## Were any barriers to delivery experienced?

No barriers to delivery were experienced as such, but some challenges were occasionally mentioned by teachers. These were related to (1) time demands, (2) focusing on certain genres of writing that lend themselves well to IPEELL, and (3) extending the more able writers, as discussed below.

### Time demands

At the very beginning of the implementation of IPEELL in the schools, some teachers mentioned the time demands that it placed upon the teacher in terms of setting it up, planning, and preparing the working wall. However, it was felt that this was time well-spent, as shown in the following comment about the value of efficient wall displays:

*'The working wall takes a lot of time. The wall at the back of the classroom in the corner—it took me a day to create it and then every time we do a new genre, the same kind of work goes up so every time there's all of the information. However, you get out what you put in and I know that the children use it and I know that all of them—when we're starting a new topic, when we're doing our writing—they will be going up to the board, they will be using it, they will be engaged with that, so the time I put into it is worth in terms of what they're getting out of it' (School 58).*

Issues with time demands were also raised with teachers having to develop a greater range of model texts in addition to the ones available through the teacher manual (as discussed in the Teacher Manual section). Furthermore, peer marking was sometimes perceived to be particularly time consuming and some schools did not make consistent use of peer marking due to curriculum time demands (as discussed in the section Quality of the IPEELL Delivery).

### Writing genres

Another challenge concerned the inappropriate use of IPEELL that was observed in two schools. Teachers very often mentioned that IPEELL worked particularly well with non-fiction writing, chronological reports, and narrative writing. In two lessons, IPEELL was being applied to poetry (School 23) and writing a play script (School 53). In the first school, the IPEELL planning frame had been adapted to the poetry genre but the main elements of the frame had been completely altered (for example, the Points part of the planning frame did not form part of the adapted one). In the second school, there was confusion about the writing genre. The teacher mentioned that a filmmaker was working with the children on their documentary writing work but this focus was lost, as more emphasis was put on writing a script and making a film which deviated from the documentary focus that could have worked well with IPEELL. Good practice was observed in another school where the teacher—realising that IPEELL did not work well with a particular long detective story that the children were writing—stopped using the IPEELL structure towards the end of that particular writing task:

*'It worked for the sections of the story that were very clearly defined, like a beginning, a character, a setting ... [but] when we put the story together, towards the end of the story it became too cumbersome and then it didn't work for me really, so we then stopped and picked up again when we did some science reports' (School 50, Spring 2016).*

### Extending the more able writers

A challenging issue that had to be addressed concerned the difficulty in finding ways to extend the more able writers who might be able to quickly 'tick the boxes' but the initial mark scheme 'didn't actually account for their style, their composition skills, which were of very high level', as one teacher said (School 50, Spring 2016). In Year 2, the CEP developer mentioned to the Year 5 teacher in one of the control schools (School 54) that the planning sheet can be divided into more than three boxes to accommodate the more able writers' desire to write more. These children could even do their own

planning from the start: ‘I would be saying to [the child], “this was an example of how to plan, you are a capable girl, you do your own planning.”’

This was an initial challenge for the schools, but many of them did find ways to help support their more able writers and this was evidenced in our later visits (for example, consulting and using the relevant resources found in the CD-ROM that was provided with the teacher manual, adapting the marking criteria accordingly, encouraging the fading out of the planning frame earlier for the more able children, and so on).

### **Is the intervention attractive to stakeholders?**

The intervention is attractive to pupils and teachers as demonstrated in the Outcomes section below where their perceptions about the strategy are presented and discussed. The following comment by a teacher in a particular school shows her enthusiasm about the IPEELL strategy which was not uncommon among teachers in the other schools visited:

*‘Writing is something that I’ve always struggled to move children on with and it’s a fantastic tool, it’s that clarity that it provides. In twenty years of teaching I’m just really impressed; really, really glad to be here at this point at this time to have this and, even if IPEELL goes away, it’s something that I will always be using’ (School 19, Summer 2016).*

## **Fidelity**

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

The intervention was delivered as intended in all of the intervention schools. The delivery was a pragmatic one, meaning that schools could use their professional judgment to apply the strategy in a sensible and realistic way based on their pupils’ needs within their particular school context.

### **What elements of the intervention are perceived to be adaptable?**

When our visit coincided with the CEP developer’s visit, we observed that a lot of support and clarification about the strategy was offered to the teachers. A common piece of advice was that the strategy should not be a straightjacket but should instead be used flexibly and adapted to the needs of the school. An example of this is the mark scheme which is flexible and should be adapted in clear and simple language to be easily understood by the children. There can be different mark schemes within a class of different abilities so that the mark scheme can be personalised for children’s different needs. The teaching would then be well aligned to the assessment criteria (assessment for learning). Another example concerns the graphing of scores which was used by all schools at the beginning of the implementation of the strategy but many stopped using it when the pupils gained more confidence in using IPEELL and becoming more independent writers. This was accepted as a positive development as teachers were reassured that graphing could stop being used when it did not serve its initial motivating purpose any more.

However, other features of the strategy, such as the writing genre, the use of mnemonics, the planning frame, aspects of self-regulation (self-talk, self-score, goal setting), motivational messages, peer/teacher scoring, and memorable experiences were non-negotiable and were expected to be evidenced in the lesson or through other sources such as wall displays, pupil exercise books, or in discussions with pupils and teachers.

## **Outcomes**

This section presents findings from our observations of the training days before discussing the perceived outcomes of the intervention for the pupils and the teachers involved.

## **Training the trainers**

A 'training the trainers' day took place in Leeds in May 2015. The aims of this training session were to familiarise the trainers who were going to deliver the IPEELL with memorable experiences intervention to teachers in Leeds and Lincolnshire with its background and aims, the writing process, its different elements and stages, and pupils' involvement in it. The trainers were given a teacher manual and CD which had been developed by the Calderdale Excellence Partnership (CEP) and which they could use in their own training with the teachers. CEP explained in detail the original efficacy trial which had been funded by the EEF, and introduced the trainers to the ongoing effectiveness trial.

### *Strengths*

The six stages of the writing cycle were introduced and each one was elaborated in detail. There was good discussion about 'memorable experiences' and the developers explained the links between memorable experiences and the writing process using IPEELL thoroughly. A discussion about writing genres took place with the developer clearly explaining the use of the mnemonics PAT and IPEELL. The importance of setting a 'cold' task was emphasised in setting a baseline against which children's writing would be compared. The use of the planning frame was thoroughly explained and other important aspects of the intervention—the development of a mark scheme, self-scoring, graphing of scores, goal setting and peer review—were also discussed in detail. The importance of positive self-talk in the form of motivational messages before, during, and after the writing task was exemplified using an analogy with sport and examples from the original project in the manual. Examples of activities that the trainers could do during their own training with the teachers were shared.

Overall, the training was logical and coherent, and covered all the main elements that were needed for implementation. The pace was appropriate and a purposeful learning environment was established.

### *Limitations and suggestions*

The training covered in detail all aspects of the intervention. The developers guided the trainers through the material, thoroughly clarified important aspects, provided plenty of examples from their own experience in the classroom but they did not model any of the activities. The trainers were advised to use the material as they saw fit and use their professional judgement to adapt the information given in their particular context. However, they were informed that some aspects of the intervention were 'non-negotiable', such as the use of the planning frame, the development of the marking scheme, scoring, graphing, goal setting, and peer review.

Due to the lack of modelling of any activities, it would be useful if, in the future roll-out of the cascading training model, the trainers could observe the developers (or other trainers) delivering the training to teachers so that they could then replicate it in their own training. The trainers were adequately informed about all the IPEELL elements and were provided with engaging examples, but observing IPEELL being delivered in a real context would be desirable for the teachers involved.

## **Training the teachers**

Two 'training the teachers' days were observed, one in Leeds (June 2015) and one in Lincolnshire (June 2015). Two members of the process evaluation team observed each training session. One of the CEP developers was also present and offered guidance and support to the trainers whenever needed. The aims of this session were for the trainers to train the teachers who were going to deliver the IPEELL with memorable experiences intervention in Leeds and Lincolnshire. The teachers were familiarised with the background and aims of IPEELL, the writing process, its different elements and stages, and pupils' involvement in it. The teachers were also given the teacher manual and CD which had been developed by the Calderdale Excellence Partnership (CEP) to help them with their IPEELL teaching in

their schools. CEP explained in detail the original efficacy trial which had been funded by the EEF, and introduced the teachers to the ongoing effectiveness trial.

Another training session delivered by the CEP developers to teachers in the Halifax local authority was also observed. Halifax LA was funding the intervention and training, and we were able to observe this ‘train the teachers’ session as a control with the observations of the ‘train the trainers’ session and the teaching sessions run by trainers trained by CEP. Two members of the evaluation team observed this session.

Our observations of the Leeds, Lincolnshire, and Halifax training days are presented in the sections that follow.

### **Leeds training day**

#### *Strengths*

The trainers briefly explained the background to the strategy and referred to the importance for children to have real experiences as a basis on which to build their writing. The teachers worked in four large groups and they were given a lot of opportunities to share ideas and interact with others in their group. There was a good level of interaction among teachers during their small group work. There was a discussion about writing genres and teachers were asked to deconstruct a particular piece of writing on to a standard planning frame. The mnemonics PAT and IPEELL were introduced and the other planning frames (discussion, descriptive, and narrative) were explained. Other aspects of the writing project—such as marking and self-scoring, graphing, goal setting and peer review—were mentioned during the training and the manual was used extensively to provide more detailed information. All aspects of the intervention were mentioned but some were discussed in more detail than others.

A particular strength of the training was the demonstration of a video clip where a ‘think aloud’ protocol was being modelled to teachers; the trainer in the video was taking the prompts from the floor and was responding. This was good modelling of a piece of writing being developed with the thought process behind it made explicit.

#### *Limitations and suggestions*

The training provided on the day was limited in terms of explicit modelling, adequate explanation, and direct interaction with the audience. Firstly, there was little explicit modelling of important aspects of the writing project, such as the use of the planning frame and positive self-talk. Secondly, the explanations given to aspects of the project, such as self-scoring, graphing of scores, peer review, and goal setting, were not fully developed. For example, the link between goal setting and assessment for learning and the development of the mark scheme were not explicitly discussed. Finally, more direct interaction with the audience in encouraging teachers to feed back to the whole group after smaller group activities and more detailed explanation of aspects of the writing strategy may have enhanced the quality of the training.

The CEP developer’s presence during the training was valuable. Input was offered to the trainers during breaks. Questions were also answered and advice was given to the teachers during discussion times. At the end of the session, feedback was given to the trainers which they took on board and, as a result, the second training day improved substantially (according to the CEP’s observation on the second training day).

*Even though the training was adequately informative and had the aforementioned strengths, it was lacking in providing breadth of practical examples and more depth to the explanations provided. We believe that this enriched approach would be derived from more practical experience and thus, seeing the intervention being delivered by an experienced trainer would help address this. On the basis of*

*these observations, we recommend that having an experienced trainer who has seen the intervention been delivered and can share practical examples with the teachers from experience would be beneficial in the future roll-out of the cascading training model. It would also be useful for the trainers to watch an experienced trainer train the teachers first.*

### **Lincolnshire training day**

#### *Strengths*

The background to the project, including the original efficacy trial which had been funded by the EEF, was clearly explained. The commitment required by each school and by CfBT (Lincolnshire Centre for British Teachers) was also clarified. The trainer talked about the importance of first hand memorable experiences and some discussion around these real experiences and the writing process using IPEELL followed. The trainer mentioned different genres and how writing in each one can be supported through four different planning frames (standard, discussion, descriptive, and narrative). The 'cold' writing task was explained clearly and it was linked to setting a baseline against which to demonstrate improvement. A discussion followed about how the task fitted with existing marking policies and whether it should be on paper or in books. CEP also added that the 'cold' task can be used as a powerful piece of evidence for moderation and it can enhance children's motivation in writing.

A model text ('Penguins', an example of a non-chronological report) was analysed in terms of purpose, structure, and language features using the questions posed by the manual (p. 2.1). The mnemonics PAT and IPEELL were clearly explained and further exemplified by applying the model text to the planning frame. A later activity based on the use of a video clip to stimulate ideas and fill the planning frame was interesting. The trainer also modelled copying a sample text she had already written but her thinking process behind the writing was not made explicit. Teachers were referred to the manual and asked to devise a mark scheme. This process was defined as setting 'success criteria with a mark'. A very good discussion around this aspect followed with reference to new assessment concepts of mastery with depth (exceeding expectations), mastery (secure at age-related expectations), and working towards mastery. Positive self-talk and the use of motivational messages were covered well with reference to the manual. Graphing of scores, goal setting, peer review, and scoring of writing were covered but not in great detail.

Overall, the materials and information were well presented, the feedback to the whole group from small group discussions was useful; there was some good modelling and a range of small group activities were planned and delivered.

#### *Limitations and suggestions*

Even though different writing genres were mentioned, this could have been done in greater depth. Greater interaction between trainers and teachers during the activities might have been more beneficial. Aspects such as the 'thinking aloud' process could have been modelled more spontaneously so that the thinking process could become evident.

The presence of the developer, CEP, during the training was useful as input and clarification was provided at times. Most importantly, particular examples were provided about how the intervention has worked in practice in the classroom.

*On the basis of these observations, we recommend that having an experienced trainer who has seen the intervention being delivered and can share practical examples with the teachers from experience would be beneficial in the future roll-out of the cascading training model. It would also be useful for the trainers to watch an experienced trainer train the teachers first.*



## What are the perceived outcomes of the intervention?

The perceived outcomes of the intervention are here discussed from the perspective of the pupils and the teachers in the schools visited.

### Pupils' attitudes

In all school visits conducted from autumn 2015 to spring 2017, pupils talked with enthusiasm about their IPEELL writing lessons. They all felt that the IPEELL writing process had helped them develop their writing and were happy to see the progress they had made between the initial cold task and their final piece of writing. Seeing the improvement in their writing was thought to be motivating for the pupils and gave some 'a boost of confidence' (School 58). Pupils were happy to see their higher marks on the hot task compared to their marks in the initial cold task. They also described how more confident they felt when they got to complete the hot task after the planning phase had taken place:

*'You are really confident to do it because in the cold task it's all from your head, you haven't planned it out but when you're doing the hot task, you've planned it, you can just look back if you are feeling stuck with your work' (School 59, Spring 2016).*

The pupils also commented highly on the opportunity they had in class to look at a model text on a particular genre and discuss it with the teacher. They also welcomed the opportunity they were sometimes given to look at 'a bad one to see what a bad one looks like to compare' (School 19, Leeds).

### IPEELL Structure and Planning

All pupils interviewed commented very highly on the IPEELL structure. Compared with their previous writing experiences, some pupils mentioned that they preferred using the IPEELL writing frame because it offered a useful structure which they could refer to while writing:

*'Cause when we used to do it, we needed to remember what we had been taught and then sometimes you would forget it. But with the IPEELL it's better because you can go back to it' (School 58, Summer 2015).*

*'It has really helped us a lot with the planning 'cause we used to just like do normal plans and it was difficult for me ... I didn't know what to write but the IPEELL structure gives you the paragraphs and what you need to write' (School 19, June 2016).*

The planning frame using the mnemonics PAT and IPEELL helped the pupils structure and develop their ideas. It provided children with the opportunity to develop ideas, include these in their planning, and then develop them further in their writing:

*'It provides you with little boxes, so you don't have to write an awful lot, it just gives you enough room to write something ... It gives you the chance to think about the introduction and the characters' (Digby 27, Spring 2016).*

*'Sometimes I struggle with writing, I can either write a short paragraph or a really, really long one. So the planning sheet is really good because it equally splits it up so there is enough room to write as much as you want to' (School 50, Spring 2016).*

Working with IPEELL also enabled the pupils to enrich their vocabulary and 'make better use of language features' which contributed to improved writing as evidenced in their higher scores in the marking sheet:

*'We are expanding our word choices in class and we are using more dictionaries, and in our marking sheet we have loads of different words to try and put in our work, so we put lots of information and that will get us more points' (School 81).*

As a result, most of the children interviewed mentioned their increased enjoyment of writing with the introduction of IPEELL in their literacy lessons as many felt that they had become more competent and confident writers, as the following quotes illustrate. One pupil, in particular, said: 'it is very helpful now because we do deeper learning' (School 48, Summer 2016).

*'When we use the IPEELL, it makes our writing make more sense. Before when we didn't use IPEELL, the sense in my writing was really down, it didn't make sense, now my work makes better sense. Now it's easier.'*

*'My writing has really improved from Year 4, before I was writing anything and it didn't make sense, now I know what to write about with the IPEELL frame.'*

*'When we were younger, we used to write anything, I used to start off and then I would write about something else' (School 33, Summer 2016).*

However, a group of high-achieving girls in one school mentioned that they did not enjoy using IPEELL as they had found it quite restrictive—they did not feel that IPEELL gave them enough freedom to plan in their own individual way.

### **Self, peer, and teacher marking**

An important part of the writing process for the pupils was the marking sheet which allowed them to self-mark and peer-mark their work with final input from the teacher. The marking sheet was, in most cases, developed by the teacher but the marking criteria were then discussed in class. However, pupils appreciated the opportunity they sometimes had to come up with ideas about marking criteria in small groups, share them with others, and see their ideas displayed in the classroom before the final list was compiled by the teacher. Some pupils felt that, as they gained more writing experience, they were in a better place to contribute to the marking criteria in order to target their own individual needs. The following comment illustrates the pupils' input in the development of the marking sheets in one particular school:

*'On some, Miss does it by herself but on some we do it together, some we did it all by ourselves—Miss gave us loads of cards with things that could go on the mark scheme and were each given a section. My group were given links and we were given a paper with links on and we had to choose the certain things and how many marks we would give to each' (School 23).*

The pupils mentioned that the habit they started getting into of marking their own work as part of IPEELL had facilitated their ability to self-evaluate and make progress with their work. The children enjoyed marking their own work and then having the chance to compare their own views with their partner's and teacher's views. They mentioned that they tried to be strict with themselves and avoid marking themselves too highly to avoid being embarrassed in case their partner or teacher gave them lower marks. The children also mentioned that their ability to mark their own and others' work improved with time, which helped them gain more awareness of their learning and level of achievement in writing. They all agreed that peer marking was helpful because it provided 'a different point of view to the teacher's' (School 33).

Moreover, self-marking was thought to be a powerful tool of the strategy because it enabled the children to see where they were good at writing and what they needed to work on. This seemed to help with feelings of frustration when the mark given previously by the teacher was lower than expected ('I used to get mad with myself, now I mark it myself and I know how I am doing and how I can improve', School

48, Summer 2016). Marking one's own work became an enjoyable process as pupils developed increased understanding about how they could improve their work. The pupils mentioned that they tried to be fair in their self-marking but were also strict with themselves.

Furthermore, peer marking seemed to work well because it allowed discussions to take place about how a piece of work could be improved both through listening to a peer's advice but also through reading somebody else's work. Peer marking could sometimes lead to some tension when lower marks were given from peers but an honest conversation about ways of improving one's work seemed to follow in most cases. As one pupil said, 'Peer marking is helpful because if you get stuck on something, if you look at someone's work, you can use bits of information that you could use in your next piece of writing' (School 22, Autumn 2016).

However, pupils also mentioned that, occasionally, peer marking was not being taken seriously and was not conducted fairly. For example, it was felt that marks or feedback given by a partner would not be honest and, therefore, marking each other's work would become a meaningless process:

*'I don't really think that peer marking is very good ... If it's your enemy, for example the person sitting next to you, they are going to give you something really bad and your friend would give you a high mark and you sometimes have to lie. So we won't really improve.'*

*'I would take it personally if you are my friend and you give me a bad score' (School 19).*

### **Graphing of scores**

Graphing was perceived to be an enjoyable part of the process as it provided a powerful but simple visual tool for observing the progress made and enhanced pupils' motivation and positive feelings about writing. This is illustrated in the quotes that follow:

*'It helps us to understand it and Miss to understand it and then we can see how much we improved' (School 60, Autumn 2015).*

*'It makes you feel good when you are colouring in your next score compared with your first one. In the first one, you might think it is not much but in the second one you might think "oh, this is a lot more"' (School 58).*

*'It's quite simple, it's not really complicated graph, it's really easy, you can understand it really easily' (School 59, Spring 2016).*

Graphing was a more regular feature of pupils' writing work in Year 5 but it was not often perceived as important in Year 6 once pupils had become more independent writers. Year 6 pupils felt that they were able to look back and compare their marks without needing the visual reference that graphing provided (for example, School 22, Autumn 2016).

### **Goal setting**

Another reason why pupils felt more confident with writing using the IPEELL planning frame was because they felt that they now knew how to improve as they had targets to work on. Pupils in different schools mentioned that there was a general goal for everyone in the class for the different writing genres and some more specific targets for individual children. The task of having to identify personal writing targets was thought to be useful. A number of pupils across the schools visited mentioned that their target-setting ability had become more accurate with time and they found that the teacher often agreed with them but they also welcomed the teacher's feedback on their writing.

Goal setting seemed to work effectively with pupils taking the initiative to set their own goals before the teacher would add to those goals or approve the child's own target. Some pupils mentioned that they enjoyed setting their own targets before the teacher's input as this gave them a chance to 'cool down and think about what they had done' (School 48, Summer 2016). Similarly, a pupil in another school said, 'basically you set your own achievement and then you achieve it' (School 53).

However, target setting seemed occasionally to be mainly teacher directed. In one school (School 19, Summer 2016) this was common practice, according to the pupils interviewed, and it was welcomed by all. These pupils thought that the teacher was the only one capable of looking at their work in detail as they did not consider themselves able to evaluate their own work and come up with an accurate judgement:

*'I just like it when the teacher gives it you 'cause you are a child, you don't know your full potential yet.'*

*'When you're marking for yourself, you think everything is right.'*

*'I used to set myself targets and I would just ignore them, it was just something I wrote down, the teacher gives more attention to your work.'*

### **Memorable experiences**

The pupils interviewed mentioned a number of interesting places they had visited. They talked with enthusiasm about how these visits facilitated their subsequent writing as they provided them with experiences that they could vividly remember. Having real experiences was perceived to facilitate pupils' writing because they could then write more realistically about true events. For an instruction-writing topic, for example, pupils had been to a nature reserve and learnt how to make a fire, which then gave them the information needed to confidently write the fire-making instructions (School 58, Summer 2015). In another school, after a visit to a local park, the pupils wrote a letter to the council to ask for improvements in the park (School 59, Spring 2016). Memorable experiences were perceived as being important in increasing pupils' knowledge and developing experiential learning which were valuable for the subsequent writing task. Pupils were then able to 'picture' the experience without having to 'make something up' (School 23). As one particular pupil said, 'You don't get the gist of it so you don't know what to write about, but if you go there, you know what to write about' (School 33).

### **IPEELL becoming embedded**

It was also mentioned in the interviews that pupils were often using the IPEELL structure in other subjects across the curriculum using the skills gained through IPEELL even when not prompted by the teacher to do so. IPEELL skills were used in writing in science, religious education, or history, for example:

*'It hasn't just helped with the writing, it has also helped in science when we have had to write in science. We are so used to writing using the IPEELL frame, so we use it when we write anything, it has helped a lot.'*

*'When we are in science and we are planning, I also use the IPEELL frame cause I just think, oh IPEELL!' (School 33).*

*'If we're told we're going to be doing any writing in other subjects, in science or history, we just instantly think about the structure ... so we go straight into IPEELL mode' (School 50, Summer 2017).*

Towards the end of the first year of the implementation of the strategy, it was observed that, in some schools, the planning frame had started fading out which was a sign that the strategy was becoming

well embedded. The children also mentioned that they could now remember the meaning and different elements of PAT and IPEELL clearly ('I remember it in my head') and that they would still be able to use it even if the words were removed from the planning frame ('I would still remember it, cause we've done it so many times', School 48, Summer 2016). In some observations, only the IPEELL initials were used and acted as a reminder of the structure. The planning was sometimes needed less and less:

*'I (still need the planning) cause sometimes I might forget, I need reminders—but not so detailed any more.'*

A girl from the same school said:

*'We have our daily logs and we usually plan them in there, we write IPEELL down on the side and we usually plan them there cause we don't need it as detailed as it was at the start' (School 11, Summer 2017).*

At the end of Year 6, pupils often mentioned that the IPEELL planning frame was used consistently from the beginning of Year 5 but, as they were so much used to it, they were much more comfortable to plan and write without the help of the planning frame. However, the pupils also mentioned that those children who were still not very comfortable could choose to continue using the IPEELL planning frame to help structure their ideas (School 50, Summer 2017).

Furthermore, when asked about their performance in cold tasks towards the end of the second year of the project, many pupils said that their cold tasks had massively improved and they were now easier. As one pupil said:

*'Before, you would have to look at everything and you would have to know what you were doing, but now you can just picture it and you can just do it without' (School 11, Summer 2017).*

The pupils interviewed in this school found the process of charting the comparisons of their marking in the hot and cold tasks motivating and enjoyed seeing the improvement they had made. Furthermore, when asked how they felt about the cold task after having been doing it for almost two years, they mentioned that they were 'more confident' because they knew what to do:

*'The first cold task, I was quite worried about it, I had a go but I was not concentrating so much, and it turned out quite good. But now, because I have been doing it for a while, I am more confident.'*

*'I just embrace it now, we've got a cold task, let's get on with it and see what I can do.'*

Last but not least, the Year 6 pupils interviewed had found their IPEELL work useful in preparation for the SATs where they 'used IPEELL to do the writing, it helped'. Some pupils also mentioned that they will retain the IPEELL structure when they move to the secondary school, probably using 'the marking scheme, some points, punctuation, links back to the introduction', hoping that writing in secondary school would become easier because of their developed IPEELL writing skills:

*'It will help us structure our writing 'cause secondary school will be harder and it can help us do the work more easily' (School 11, Summer 2017).*

### **Teachers' views**

In the schools visited, all teachers felt that IPEELL had been particularly beneficial for helping pupils to structure their writing work, and worked well with all ability levels. The simplicity of the structure and its application to different writing genres was perceived to be key to its success. In addition, the planning sheet, which could be structured slightly differently for the different ability levels and the various writing

genres, was very useful as it acted as a 'sort of comfort blanket' for the children. As a result, children seemed to have a better understanding of what a good structure in writing looks like, as the following example illustrates:

*'So they have more of an understanding of what goes where, whereas before we had different headings for everything, for every different type of writing, so I think they can see the common threads within different genres of writing more now than before ... that's something we were looking to improve, especially for lower ability children to help them with that structure and that sort of formalised lesson, give them a framework really' (School 50, Spring 2016).*

Boys, in particular, were seen to be benefitting a lot from IPEELL, as the Year 6 teacher reported in the interview, and were shown to make the most progress because of the IPEELL structure and their preference for non-fiction writing which was well-aligned with IPEELL (School 22, Autumn 2016). Similarly, teachers thought children of lower or middle ability found the IPEELL structure very useful:

*'Some would ramble on and wouldn't be precise enough, the structure helped them to be more precise. They can all achieve a piece of sustained writing now, it's the quality that differentiates the different children' (School 22, Autumn 2016).*

It was evident, in a number of schools, that the children were paying particular attention to the content and structure of their writing work, the links and the language used as a result of IPEELL instruction. In particular, some lower and middle ability children ('children that are top of your bottom and bottom of your middle', School 50) started writing more lengthy pieces because of the 'concrete structure' that they were provided with as part of IPEELL. It was felt that those lower ability children that would normally require more help than the others did not need as much scaffolding and support as they got more experience in writing using IPEELL. Consequently, a Year 5 teacher had noticed that, when mixed-ability groups were used in the literacy classroom, 'that bottom group is shrinking' (School 50, Spring 2016).

As teachers mentioned in another school visit (School 11, Summer 2016), children of lower and middle ability became more confident 'in the mechanics of it' and they were then freer to work better on the content and the quality of their writing:

*'I would say it has worked for everybody, probably the best for the lower to the middle ability pupils. It has given them a structure to hang things on. The others are fine and they've learnt from it but they've moved on from such a structured approach, I would say. So, everyone has made progress, I would say, but particularly those lower end learners' (Year 6 teacher).*

*'It's more the fact that they can now structure it and they are confident with what they need to put where and I feel that's really helped the lower and the middles, who before didn't really know how to structure it but now they can, it's really brought them on and they're confident to write, and they're writing a lot more as well which is nice' (Year 5 teacher).*

Furthermore, one teacher mentioned that children who have English as an Additional Language in her school could make impressive use of technical vocabulary as a result of IPEELL. This teacher attributed the children's language development to their IPEELL writing and felt that the children's writing had 'drastically improved':

*'I've got children in there who have been learning English for six months and are using words like bacteria and fungi in their writing; I am really impressed' (School 59, Spring 2016).*

Higher ability children also benefitted from IPEELL. As is shown in the following teacher's comment, the higher ability children would develop their plans more quickly and in more detail while the children of lower ability would put less on the plan but were still able to access it:

*'I think it has been quite accessible for all really. Obviously, the higher ability got their heads round it a lot quicker and they were then able to use the plans quicker. I think the plan was the bit that took them the longest to get, the breaking down between the IPEELL and what to put where, so I think in my case my higher ones have grasped that quicker and put more detail on the plans and helped themselves with their plans more, and the lower ability put less on, but they can still all do it, they can still access it, mine were all fine' (School 19, Summer 2016).*

As IPEELL could be adapted to the different pupils' needs, it seemed to work well with all ability levels. However, some challenges encountered with extending the more able writers further were mentioned by some teachers and these are discussed in the section that addresses the barriers to delivery.

Another important outcome of IPEELL, according to the teachers interviewed, was the children's developing ability to use the IPEELL structure across the curriculum (for example, in topic work) showing a possible transference of learning in other pieces of writing outside of the literacy class. This was raised in a number of schools. In the following example, teachers noted that pupils were using elements of IPEELL in all of their writing work across the school. The quality of the writing was reinforced across all subjects and the children seemed to be thinking in terms of the IPEELL structure in their overall writing work:

*'Yes, 'cause they always say five paragraphs 'cause I need to ... know how to structure it, and even in topic or science or RE or anything they've got, they'll say "I need to do my IPEELL" and they know how to set it out' (School 11, June 2016).*

In another school, the IPEELL strategy seemed to be well embedded as it was used spontaneously across the school in all of the children's writing work. According to the teacher:

*'We apply it to our topic writing, we use the exact same process, we use the mark schemes in topic writing so that's been really powerful. As a school, we needed to work on our cross-curricular links. I think they use it in their RE writing, they use it, basically, in any kind of topic that they do writing for without being prompted. I can see them writing IPEELL in the margin. That's really good, because topic writing again tends to always be not as good as their literacy writing because they're like "oh it's topic you know we don't need to try as hard", but because they've had the mark scheme in their head it's like "oh this is like a literacy lesson so I've got to produce a good quality piece of writing"' (School 23, Summer 2016).*

Teacher views were also heard about how the standard of writing had improved across the curriculum and that the children had been observed increasingly using the IPEELL structure in a spontaneous way when writing cold tasks.

Finally, it was a common view among the teachers interviewed that the children were gradually developing increased self-regulation, evidenced in their ability to be self-reflective in marking, setting goals, and editing their work. Across all schools, children had become better able to evaluate their writing and provide more accurate targets with time. It was often mentioned that the constant comparison between initial writing tasks and writing that was produced after the IPEELL planning enabled children to identify the progress they were making. As an example, the teacher in a particular school mentioned that the Year 6 class had really taken IPEELL on board by editing and marking their work in an engaging way as well as setting their own targets. The children in this particular class were actively involved in the marking process by being encouraged to jointly write a mark scheme with the teacher. This, together with regular IPEELL writing work, contributed to the children developing a very good understanding of the IPEELL process:

*'They're setting their own goals, they're target-setting, they're going back and editing and proofreading their work, they've set the mark scheme—we wrote a marks scheme together but I didn't write it for them, they actually created it with me ... So this class have really rolled with it but I think because I went on the training and I had a really good understanding of it, now that I'm at the end of*

*the year we go over the same process every single time for every single text, so even the kids kind of anticipate what's going to happen now' (School 23, Summer 2016).*

### **Were there any unintended consequences or negative effects?**

No unintended consequences or negative effects were found from the process evaluation.

### **Formative findings**

#### **Are there any ways that the intervention can be improved?**

On the basis of the process evaluation findings, we recommend the following ways in which the intervention can be improved:

1. Regarding the training of teachers: having an experienced trainer who has seen the intervention being delivered and can share practical examples with the teachers, from experience, would be beneficial in the future roll-out of the cascading training model. It would also be useful for the trainers to watch an experienced trainer train the teachers first.
2. Teachers delivering the IPEELL strategy should be encouraged to fully understand, from the start, that:
  - A. Some features of the strategy are non-negotiable and are expected to be evidenced in pupils' IPEELL work. These features concern a focus on an appropriate template for IPEELL writing genre, the use of mnemonics, the good use of the planning frame, aspects of self-regulation (self-talk, self-score, goal setting), motivational messages, and peer/teacher scoring.
  - B. IPEELL should not act as a 'straightjacket' but should instead be used flexibly and adapted within the context of each particular school to help all learners make good progress and reach their full potential.



## Conclusion

### Key conclusions

1. Pupils who used IPEELL for two years made around two months' additional progress in writing, compared to pupils who did not. This result has a high security rating. The result was similar when considering only pupils eligible for free school meals who used IPEELL, who also made around two months' additional progress.
2. The evaluation also measured writing results for pupils who used IPEELL for only one year, using a different writing test. This test showed that after one year, IPEELL pupils had made *less* progress than comparison pupils. This result had a very high security rating. Again, the results were similar for pupils eligible for free school meals.
3. These writing results are not statistically significant. This means that even if the intervention had no impact, the probability that, in this trial, we would have observed an effect size as large as the one found is greater than 5%.
4. Pupils who used IPEELL for two years made less progress in reading, spelling, and maths compared to pupils who did not. It is possible that an increased focus on writing led to less time spent on these topics.
5. The training of teachers was lacking in both the provision a breadth of practical examples and in fully modelling some aspects of the approach. This may have been because the training in this trial was carried out by inexperienced trainers and it may, therefore, be useful for the trainers to watch an experienced trainer train the teachers first.

### Interpretation

These trials (Trial 1 – a one year test of the approach and Trial 2 – a two year test of the approach) evaluated an intervention adapted from an intervention previously developed in the U.S. IPEELL was developed by the Calderdale Excellence Partnership and is based on the U.S. Self-regulated Strategy Development (SRSD) intervention, combining this with memorable experiences, which pupils can use as a stimulus for their writing. Previous U.S. evaluations of SRSD found positive effects for the intervention in individual trials and systematic reviews of the trials. Most of the research was undertaken by the developers of the intervention, however. A previous U.K.-based efficacy trial of IPEELL undertaken by some members of this evaluation team (Torgerson et al., 2014) found a large positive effect for the intervention (0.7). The effectiveness trials reported here differ from the efficacy trial in a number of ways: the intervention was specifically developed to be more acceptable to U.K. schools; the training in the intervention used a cascade or 'training the trainers' model; and independent testing for the primary outcome measure was undertaken for Trial 2. The observed outcomes for these two trials did not find consistent positive effects, which contradict the previous evidence in the literature. The results from the two trials reported here may differ from the results of the efficacy trial due to a number of factors: the professional development model (training the trainers) may have been less effective in training the teachers in the intervention; the outcome measure may have been less susceptible to the introduction of potential bias; and the implementation of the intervention at scale may have been less optimal. The results may differ from the results of the U.S.-based evaluations due to additional factors such as the addition of the memorable experiences component which is not part of the SRSD intervention, and the developer-led nature of the U.S. evaluations.

### Limitations

The study had lower statistical power than expected, which may have meant, particularly for Trial 2, that an educationally important difference of IPEELL could have been missed.

The results can be generalized to other primary schools with similar demographics to the schools (in Leeds and Lincolnshire) in this evaluation, using the same manualized intervention delivered through the cascade model of training and including memorable experiences. However, it is not possible to disentangle results from IPPEEL and memorable experiences; in other words, the extent to which the two components impacted on the results cannot be known from the design of this trial.

### **Future research and publications**

Future research questions that need answering include the effectiveness of IPPEELL as an isolated intervention (that is, without the possible confounding of the memorable experiences). This needs to be tested, although it seems unlikely that the memorable experiences had this impact. However, it would have cost implications. Given the better outcomes in Trial 2 compared with Trial 1, IPEELL might be more effective when started among younger children.

## References

- Aerts, M., Geys, H., Molenberghs, G. and Ryan, L. M. (2002) *Topics in Modelling Clustered Data*, Chapman and Hall/CRC.
- Altman, D. G., Moher, D. and Schulz K. F. (2012) 'Improving the reporting of randomised trials: the CONSORT Statement and beyond', *Statist. Med.*, 31, pp. 2985–2997.
- Andrews, R., Torgerson, C., Low, G., McGuinn, N. and Robinson, A. (2006) 'Teaching argumentative non-fiction writing to 7–14 year olds: a systematic review of the evidence of successful practice', technical report in *Research Evidence in Education Library*, London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T. and Rothstein, H.R. (2009) *Introduction to Meta-analysis*, John Wiley and Sons. DOI: 10.1002/9780470743386
- Cook, T. D. and Campbell D. (1969) *Quasi-experimentation: Design and Analysis Issues for Field Settings*, Boston: Houghton Mifflin.
- De La Paz, S. and Graham, S. (2001) 'Explicitly teaching strategies, skills and knowledge: Writing instruction in middle school classrooms', *Journal of Educational Psychology*, 94, pp. 687–98.
- Department for Education (2013a) 'National Curriculum Assessments at Key Stage 2 in England, 2013' (Revised), (Reference IdL SFR51/2013). Statistical First Release.
- Department for Education (2014) 'National Curriculum Assessments at Key Stage 2 in England, 2014' (Revised), (Reference IdL SFR50/2014). Statistical First Release.
- Hedges, L. V. (2007) 'Effect sizes in Cluster-Randomized Designs', *Journal of Educational and Behavioural Statistics*, 32, pp. 341–70.
- Shadish, W.R., Cook, T. D. and Campbell, D. T. (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston, MA: Houghton, Mifflin and Company.
- Standards and Testing Agency (2016) '2017 national curriculum assessments Key stage 2 Interim teacher assessment frameworks at the end of key stage 2 (July 2016), Crown Copyright: available online [here](#).
- Torgerson, D., Torgerson, C., Ainsworth, H., Buckley, H., Heaps, C., Hewitt, C. and Mitchell, N. (2014) 'Improving Writing Quality: Evaluation Report and Executive Summary' (May 2014), Educational Endowment Foundation (EEF): available online [here](#).
- Torgerson, D. J. and Torgerson, C. J. (2008) 'Designing Randomised Trials in Health, Education and the Social Sciences: An Introduction', Palgrave Macmillan.
- Verbeke, G. and Molenberghs, G. (2000) *Linear mixed models for longitudinal data*, New York: Springer-Verlag.

## 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
£ £ £ £ £	<i>Very low:</i> less than £80 per pupil per year.
£ £ £ £ £	<i>Low:</i> up to about £200 per pupil per year.
£ £ £ £ £	<i>Moderate:</i> up to about £700 per pupil per year.
£ £ £ £ £	<i>High:</i> up to £1,200 per pupil per year.
£ £ £ £ £	<i>Very high:</i> over £1,200 per pupil per year.

## Appendix B: Security classification of trial findings

### Trial 1- All pupils

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%	5		5
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 [ ]	
3	Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%		Adjustment for threats to internal validity [ ]	
2	Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%			
1	Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	41-50%			
0	No comparator	MDES > 0.6	>50%			

- **Initial padlock score:** This was a well-conducted trial with adequate power to detect an effect size of 0.2. Attrition was at 8% based on all pupils for whom we have baseline data. One school dropped out without providing baseline data. Using the median number of pupils per school (calculated as 29 in Trial 1) as an approximation for the number of missing pupils in the school that dropped out does not cause the attrition to go above the 10% threshold = 5
- **Reason for adjustment for balance** (if made): n/a
- **Reason for adjustment for threats to validity** (if made):
- **Final padlock score:** 5

## Trial 2 – All pupils

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%	4	Adjustment for Balance [ ]	4
3	Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%		Adjustment for threats to internal validity [ ]	
2	Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%			
1	Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	41-50%			
0	No comparator	MDES > 0.6	>50%			

- **Initial padlock score:** This was a well-conducted trial with adequate power to detect an effect size of 0.2. However, attrition from the post-test was 20.1%. One school dropped out without providing baseline data. Using the median number of pupils per school (calculated as 30 in Trial 1) as an approximation for the number of missing pupils in the school that dropped out does not cause the attrition to go above the 20% threshold = 4
- **Reason for adjustment for balance** (if made): n/a
- **Reason for adjustment for threats to validity** (if made): n/a
- **Final padlock score:** 4

## Trial 1 – FSM pupils

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%	4	Adjustment for Balance [ ]	4
3	Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%		Adjustment for threats to internal validity [ ]	
2	Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%			
1	Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	41-50%			
0	No comparator	MDES > 0.6	>50%			

- **Initial padlock score:** This was a well-conducted trial, but only with enough power to detect an effect size of 0.22. Attrition from the post-test for FSM pupils was 7.7%. This figure does not include the pupils from the school that dropped out without providing baseline data. Without knowing more about the percentage of FSM pupils in this school we are unable to get a true indication of the number of missing pupils and would not be confident about calculating this number. Therefore, we have not taken the missing pupils from this school into consideration for the padlocks = 4
- **Reason for adjustment for balance** (if made): n/a
- **Reason for adjustment for threats to validity** (if made): n/a
- **Final padlock score:** initial score adjusted for balance and internal validity = 5

## Trial 2 – FSM pupils

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		
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 [ ]	
3	Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%		Adjustment for threats to internal validity [ ]	3
2	Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%			
1	Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	41-50%			
0	No comparator	MDES > 0.6	>50%			

- **Initial padlock score:** This was a well-conducted trial with adequate power to detect an effect size of 0.20. Attrition from the post-test for FSM pupils was 22.6%. This figure does not include the pupils from the school that dropped out without providing baseline data. Without knowing more about the percentage of FSM pupils in this school we are unable to get a true indication of the number of missing pupils and would not be confident about calculating this number. Therefore, we have not taken the missing pupils from this school into consideration for the padlocks = 3
- **Reason for adjustment for balance** (if made): n/a
- **Reason for adjustment for threats to validity** (if made): n/a
- **Final padlock score:** initial score adjusted for balance and internal validity = 3



## Appendix C: School Expression of Interest Form



Calderdale Excellence Partnership

### IMPROVING WRITING QUALITY EEF FUNDED PROJECT

Thank you for attending today's information session on the Self Regulated Strategy Development project.

If your school is now firmly interested in being involved in the project, please provide us with your details so that we can contact you to provide further information in due course. Schools will be invited to attend a meeting with the Project Evaluators in the spring term.

**School name:** .....

**Headteacher's name:** .....

**Contact name and position (if different to Headteacher's name):** .....

**Contact email:** .....

**Approximate numbers:**

**pupil**

Number of Y5 pupils:	Number of Y6 pupils:
..... ....	.....
Number of Y5 FSM pupils:	Number of Y6 FSM pupils:
..... ....	.....

Thank you.

Calderdale Excellence Partnership Limited, Room 121, E Mill, Dean Clough, Halifax, HX3 5AX

Telephone: 01422 255006 Email: **office@hxec.co.uk** Fax **01422 255007**

Company Number: 07703965

## Appendix D: School Agreement to participate



### Agreement to participate in the Evaluation of SRSD: The Writing Strategy

Please sign both copies, retaining one and returning the second copy to Amanda Rawson at Calderdale Excellence Partnership, Room 121, E Mill, Dean Clough, Halifax, HX3 5AX

School Name: \_\_\_\_\_

#### ***Aims of the Evaluation***

The aim of this project is to evaluate the impact on children's writing skills and ultimately their results at KS2 of Self-Regulated Strategy Development (SRSD) with memorable experiences. The results of the research will contribute to our understanding of what works in raising pupils attainment and will be widely disseminated to schools in England. Ultimately we hope that the evaluation will equip school staff with the knowledge to better support children with developing their writing skills.

#### ***The Project***

Teachers in Years 5 and 6 in schools in the intervention group will be expected to use SRSD strategies in teaching writing, with support from the Literacy Co-ordinator and the Project Team, from September 2015. Prior to that, in June 2015, there will be a training day for the above staff on SRSD strategies delivered by local authority advisers, with support from members of the Project Team. Training materials will be provided on the day to take back into school. The minimum expectation is that writing will be taught using SRSD strategies in Years 5 and 6 in 2015/16 and in Year 6 in 2016/17.

#### ***Structure of the Evaluation***

The evaluation is being conducted by Durham University and the University of York (Evaluation Team). Schools who agree to take part are randomly allocated to either the intervention group or a control group.

- Schools in the intervention group will receive training in SRSD from the Project Team during June 2015. Schools will be supported to deliver SRSD to Year 5 and 6 in 2015/2016 and Year 6 in 2016/2017 and will receive funding for memorable experiences. Schools will also receive £750 as a thank you payment.
- Schools in the control group will be asked to continue with usual teaching with Year 5 and 6 in 2015/2016. From June 2016 schools will receive training in SRSD from the Project Team. Schools will be supported to deliver SRSD to Year 5 in 2016/2017 (and Year 4 if schools wish to) and will receive funding for memorable experiences. **Schools in the control group are explicitly asked**

**not to deliver SRSD to Year 6 in 2016/2017.** Schools will also receive £750 as a thank you payment.

Random allocation is essential to the evaluation as it is the best way of investigating what effect SRSD with memorable experiences has on children's attainment. It is important that schools understand and consent to this process.

In order to find out how the intervention is working in schools we will visit a sample of schools which agree to a visit and observe some lessons and talk with the teachers and pupils in interviews and focus groups. Informed consent will be sought before we make any observations or conduct any interviews/focus groups.

The evaluation team will use school and pupil information provided by schools including KS2 results, and information from the National Pupil Database to assess any impact of SRSD: The Writing Strategy on attainment.

### ***Use of Data***

All pupil data will be treated with the strictest confidence. Named data will be matched with the National Pupil Database and shared with the Project Team - Calderdale Excellence Partnership, the Evaluation Team – Durham University and the University of York, the Department for Education, EEF, EEF's data contractor FFT Education and in an anonymised form to the UK Data Archive. No individual school or pupil will be identified in any report arising from the research.

### ***Responsibilities***

The PROJECT TEAM will:

- Deliver four training sessions (two per LA) in 2015 and in 2016, complete with training material – supply cover funding will be provided.
- Provide funding to intervention schools in 2015 to support memorable experiences in Year 5 and 6
- Provide funding to control schools in 2016 to support memorable experiences in Year 4 and 5
- Be the first point of contact for any questions about the evaluation
- Provide on-going support to the school, in partnership with the LA

The EVALUATION TEAM will:

- Conduct the random allocation.
- Collect and analyse all the data from the project
- Disseminate research findings

The SCHOOL will:

- Ensure the shared understanding and support of all school staff for the project and personnel involved.
- At the beginning of the evaluation provide school level information.
- Inform all parents of pupils who will be in Year 5 and Year 6 in the 2015/2016 academic year about the evaluation and collect in any Opt Out Forms returned by parents.

- Be a point of contact for parents / carers seeking more information on the project.
- Provide information on all participating pupils (excluding any pupils for whom Opt Out Forms have been returned) including name, UPN, DOB, FSM status, PP status, KS1 and KS2 results.
- Consent to random allocation and commit to the outcome, implementing interventions as requested.
- When requested release Year 5 teachers, Year 6 teachers and the Literacy Co-ordinator so they can attend the training day (Control schools, Year 4 teachers, Year 5 teachers and the Literacy Co-ordinator ).
- Consider talking to the evaluation team further about the project and facilitating a visit to the school from the evaluation team.

**We commit to the Evaluation of SRSD: The Writing Strategy as detailed above**

School name: \_\_\_\_\_

Head teacher name: \_\_\_\_\_

Head teacher Signature: \_\_\_\_\_ Date: \_\_\_\_\_

Head teacher Email address: \_\_\_\_\_

School Contact (if different from head teacher): \_\_\_\_\_

School Contact email address: \_\_\_\_\_

School Tel no: \_\_\_\_\_

Thank you for agreeing to take part in this research. Please return this form at the information meeting or afterwards by post to:

Calderdale Excellence Partnership, Room 121, E Mill, Dean Clough, Halifax, HX3 5AX



## Appendix E: Information Letter to Parents/Carers and Opt Out Form



[INSERT DATE]

[INSERT SCHOOL NAME]

Dear Parent / Carer

Your child's school is taking part in the *Improving Writing Quality* programme evaluation. Durham University and the University of York have been asked by the Education Endowment Foundation (an organisation funding research into education) to independently evaluate the *Improving Writing Quality* programme.

The *Improving Writing Quality* programme is being led by the Calderdale Excellence Partnership. It is designed to improve children's writing skills before their Year 6 KS 2 assessments. Good writing skills are important for all children.

To find out how well the *Improving Writing Quality* programme works some schools will use the Improving Writing Quality programme this year and some schools will not. This is decided randomly by a computer. (However all schools will continue to teach children writing skills.) Researchers will then compare results from schools that have used the programme with schools that have not. In order to do this we would like to collect information about your child from your child's primary school.

For the purpose of research, information provided by your child's school (including your child's name, date of birth, gender, free school meal entitlement, unique pupil number, pupil premium status and Key Stage 1 and Key Stage 2 results) will be linked with information about your child from the National Pupil Database (held by the Department for Education) and shared with the Calderdale Excellence Partnership, Durham University and the University of York, the Department for Education, EEF, EEF's data contractor FFT Education and in an anonymised form to the UK Data Archive. Your child's data will be treated with the strictest confidence. We will not use your child's name or the name of the school in any report arising from the research.

**If you are happy for information about your child to be used in the evaluation you do not need to do anything. Thank you for your help with this evaluation.**

If you would rather your child's school did not share information about your child for use in this evaluation please complete the enclosed form and return it to your child's school by [INSERT DATE].

If you would like further information about the evaluation please contact Kerry Bell the Evaluation Coordinator: [HYPERLINK "mailto:"Kerry.Bell@york.ac.uk](mailto:Kerry.Bell@york.ac.uk); 01904 321325 <mailto:>

Yours faithfully

Professor David Torgerson (University of York)

Professor Carole Torgerson (Durham University)

Education Endowment Foundation

Calderdale Excellence Partnership



**Improving Writing Quality Evaluation: Opt Out Form**

If you **DO NOT** want information about your child to be shared for use in the Improving Writing Quality evaluation, please return this form to your child’s school by **[INSERT DATE]**.

I **DO NOT** want information about my child to be shared for use in the Improving Writing Quality evaluation

Parent/Carer Signature.....  
Date.....

Child’s Name.....  
.....

Child’s School.....  
.....



## Appendix F: Supplementary School agreement to participate following changes to the primary outcome



### Agreement to participate in the Evaluation of SRSD / IPEELL: The Writing Strategy

Dear **SCHOOL CONTACT NAME**,

Thank you for your continued support of the independent evaluation of Self-Regulated Strategy Development (SRSD) / IPEELL, which is being delivered in your school in association with the Calderdale Excellence Partnership.

As you will be aware, changes were implemented in the academic year 2015/2016 to the way in which Key Stage 2 English assessments are recorded. As a consequence of these changes, the information that we are able to obtain from the Department for Education is no longer sufficient for us to fully evaluate the programme. Consequently, we are asking all schools taking part in the evaluation to facilitate a separate assessment for all Year 6 children in the second half of the summer term 2017. We anticipate that this assessment will take no longer than one hour and fifteen minutes.

The administration will be undertaken by the National Foundation for Educational Research (NFER), which has agreed to deliver this assessment. A representative from NFER will visit your school to administer the assessment; they will provide all papers and materials which will be taken away for marking upon completion, so that there is minimal administration for the school. NFER administrators are all qualified teachers with current enhanced DBS checks. As qualified teachers, they are experienced in the school and classroom settings and realise the importance of supporting the school and the pupils through the process. NFER will contact the school very soon to arrange a suitable time for the visit to take place between 12<sup>th</sup> June and 30<sup>th</sup> June 2017.

As this is a change to our original plan, we need to give parents another opportunity to opt their child out of the evaluation. Although the child would still complete the assessment, their data would not be shared or included in the evaluation. Consequently, we will be asking schools to circulate a new letter to parents which will include an opt-out slip. A photocopy of all opt-out slips should be emailed to Kerry.Bell@york.ac.uk. Original opt-out slips should be retained at the school.

With the exception of this change, all school responsibilities noted in the original agreement will remain the same. I have enclosed a blank copy of the original agreement for your information. I would be grateful if you could complete the attached form and return to Amanda Rawson by **XXXXXX** to confirm that you will facilitate the assessment and to provide a named contact person to liaise with NFER.

We would like to thank you again for your support and we will be issuing a thank you payment of £750 to all schools that have participated in the evaluation of IPEELL at the end of the school year.

Should you have any queries about the intervention, please contact the Calderdale Excellence Partnership at **office@hxec.co.uk** . Should you have any queries about the evaluation, please contact the evaluation coordinator Dr Kerry Bell at **Kerry.Bell@york.ac.uk**.

Kind regards,

Education Endowment Foundation

Prof Carole Torgerson (Durham University)

Prof David Torgerson (University of York)

Calderdale Excellence Partnership



Evidence for  
Excellence in  
Education

Calderdale Excellence Partnership

## Agreement to participate in the Evaluation of SRSD / IPEELL: The Writing Strategy

**We agree to facilitate the separate assessment for the evaluation of 'SRSD: The Writing Strategy'**

School name: \_\_\_\_\_

Head teacher name: \_\_\_\_\_

Head teacher Signature: \_\_\_\_\_ Date: \_\_\_\_\_

Head teacher Email address: \_\_\_\_\_

School Contact for NFER correspondence: \_\_\_\_\_

School Contact email address: \_\_\_\_\_

Thank you for continued support of this research. Please return this form by post to Amanda Rawson at:

**Calderdale Excellence Partnership, Room 121, E Mill, Dean Clough, Halifax, HX3 5AX**

## Appendix G: Supplementary parent/carer letter following changes to the primary outcome



[INSERT DATE]

[INSERT SCHOOL NAME]

Dear Parent / Carer

We contacted parents/carers in April 2015 to inform you that your child's school is taking part in the *Improving Writing Quality* programme evaluation. Durham University and the University of York have been asked by the Education Endowment Foundation (an organisation funding research into education) to independently evaluate the *Improving Writing Quality* programme.

The *Improving Writing Quality* programme is being led by the Calderdale Excellence Partnership. It is designed to improve children's writing skills before their Year 6 KS 2 assessments. Good writing skills are important for all children.

To find out how well the *Improving Writing Quality* programme works some schools are using the Improving Writing Quality programme this year and some schools are not. This was decided randomly by a computer. (However all schools are continuing to teach children writing skills.) Researchers will then compare results from schools that have used the programme with schools that have not. In order to do this we would like to collect information about your child from your child's primary school.

For the purpose of research, information provided by your child's school (including your child's name, date of birth, gender, unique pupil number and test scores) will be linked with information about your child from the National Pupil Database (held by the Department for Education) and shared with the Calderdale Excellence Partnership, Durham University and the University of York, the National Foundation for Educational Research (NFER), the Department for Education, EEF, EEF's data contractor FFT Education and in an anonymised form to the UK Data Archive. Your child's data will be treated with the strictest confidence. We will not use your child's name or the name of the school in any report arising from the research.

**If you are still happy for information about your child to be used in the evaluation you do not need to do anything. Thank you for your help with this evaluation.**

If you would rather your child's school did not share information about your child for use in this evaluation please complete the enclosed form and return it to your child's school by [INSERT DATE]. (If parents opted out previously there is no need to opt out again).

If you would like further information about the evaluation please contact Dr Kerry Bell the Evaluation Coordinator: [HYPERLINK "mailto:"Kerry.Bell@york.ac.uk](mailto:Kerry.Bell@york.ac.uk); 01904 321325 [mailto:](mailto:Kerry.Bell@york.ac.uk)

Yours faithfully

Professor David Torgerson (University of York)

Professor Carole Torgerson (Durham University)

Education Endowment Foundation

Calderdale Excellence Partnership



### Improving Writing Quality Evaluation: Opt Out Form

If you **DO NOT** want information about your child to be shared for use in the Improving Writing Quality evaluation, please return this form to your child's school by **[INSERT DATE]**.

I **DO NOT** want information about my child to be shared for use in the Improving Writing Quality evaluation

Parent/Carer                      Signature.....  
Date.....

Child's  
Name.....  
.....

Child's  
School.....  
.....

## Appendix H: Trial 2 Primary Outcome Measure

**Writing Assessment**  
**Shorter Task**  
**and**  
**Longer Task Answer Pages**

**Shorter task:** Your teacher will read through this section with you.  
You will have **20 minutes** to write your shorter piece of writing in this booklet.

**Longer task:** Your teacher will tell you when to start the longer piece of writing.

55010/EECT

CONFIDENTIAL

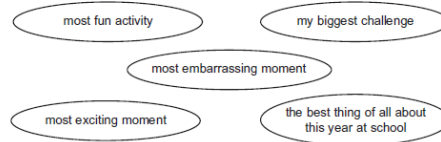
WAST

### Memories of the School Year

Imagine that your class are making a book of memories of the year at school. A copy of the book will be given to every pupil in the class.

Everyone in the class will write a page of the book about their memories.

Here are some of the things that pupils may include:



**Your task is to write your page of the book.**

You can write about real or imaginary things, or a mixture.

You can use the ideas above or think of your own topics.

#### PLANNING

Make a note of the memories you will write about.



Memories of the School Year

Lined writing area for page 3.

This is for your longer piece of writing.

Lined writing area for page 4.

CONFIDENTIAL

3

CONFIDENTIAL

4

Lined writing area for page 5.

[BLANK]

If you need more paper, ask your teacher.

CONFIDENTIAL

5

CONFIDENTIAL

6

For NFER use only.

Shorter task	Marks
ST Sentence structure, punctuation and text organisation (4)	
ST Composition and effect (8)	
Longer task	Marks
LT Sentence structure and punctuation (8)	
LT Text structure and organisation (8)	
LT Composition and effect (12)	
LT Handwriting (3)	

[BLANK]

NFER research trial booklet

Contains writing tasks © QCA 2004 and 2008; reproduced for the purpose of research.

CONFIDENTIAL

7

CONFIDENTIAL



**Writing Assessment**  
**Longer Task:**  
**Instructions and Planning**

**Your teacher will read through this booklet with you.**

You will have **45 minutes** for your longer piece of writing, including up to **10 minutes'** planning time. You may start your writing as soon as you have finished planning.

55010/EECT

CONFIDENTIAL

WALT

## Time for a Change?

Your school is thinking about making some changes to the school day.

This is an outline of the new timetable which might be introduced:

- |           |                           |
|-----------|---------------------------|
| • 7.00am  | Registration and assembly |
| • 7.20am  | Exercise session          |
| • 8.00am  | Breakfast                 |
| • 8.30am  | Lessons                   |
| • 12.30pm | Sport or homework club    |
| • 1.30pm  | School closes             |

Here are some reactions to the new timetable:

*I don't like the idea of having all the lessons in the morning.*

*I think it would be healthy to start the day with some exercise.*

*It would mean getting up very early!*

Pupils have been asked to give their views in assembly.

**Your task is to write down what you will say about this new timetable in the assembly.**

CONFIDENTIAL

2

---

## Planning

---

Reasons for your views:

<ul style="list-style-type: none"><li>•</li><li>•</li><li>•</li><li>•</li></ul>
---

NFER research trial booklet  
Contains writing task © QCA 2004; reproduced for the purpose of research.

CONFIDENTIAL

3

CONFIDENTIAL

## Appendix I: Sampling framework for school visits & schools visited

2015-2016 (proposed visits)						Schools visited		
				Leeds	Lincolnshire*		Leeds	Lincolnshire
<b><u>Intervention schools</u></b>	<b>Autumn term</b>			1	1	<b>Autumn term</b>	School 60	School 58
		School size	Small	1	1	<b>Spring term</b>	School 59	School 27
			Large	1	1			School 50
	<b>Spring / Summer term</b>	Class consistency	Separate year groups	1	1	<b>Summer term</b>	School 11	School 81
			Mixed classes	1	1		School 19	School 31 SCHOOL WITHDRAWN
		Training days attended	First training day	1	1		School 23	School 53
			Second training day	1	1		School 33	

							School 48	
2016-2017								
<b><u>Control schools</u></b>				2	2		School 86	School 52  School 54
<b><u>Intervention schools (2015-17)</u></b>				2	2	<b><i>Autumn term</i></b>	School 22	School 82
						<b><i>Spring /Summer term</i></b>	School 11	School 50

\*About one third of schools visited in Lincolnshire belonged to a federation.

NB: Some schools met more than one characteristic, but we ensured that we visited at least the stated number of schools with the desired characteristics.

### Characteristics of the schools visited

	School size	Class make up	Training day
<b>Leeds schools:</b>			
School 60	NOR 313	1xY6, 1xY6/Y5 and 1xY5	26 <sup>th</sup> June
School 59	NOR 179	1x Y6 and 1xY5	26 <sup>th</sup> June
School 11	NOR 397	2x Y6 and 2x Y5 classes	25 <sup>th</sup> June
School 19	NOR 397	1x Y6, 2x Y5	25 <sup>th</sup> June
School 23	NOR 196	1x Y6 and 1x Y5	25 <sup>th</sup> June
School 33	NOR 202	1x Y6 and 1xY5	25 <sup>th</sup> June
School 48	NOR 481	2x Y6 and 2xY5	25 <sup>th</sup> June
School 22	NOR 423	2x Y6 and 2xY5	25 <sup>th</sup> June Y6 teachers 26 <sup>th</sup> June Y5 teachers
<b>Lincolnshire schools:</b>			
School 58	NOR 106	1x Y5/Y6	3 <sup>rd</sup> July
School 27	NOR 90	1x Y5/Y6	3 <sup>rd</sup> July
School 50	NOR 236	2x Y5 and 2xY6	3 <sup>rd</sup> July

School 81	NOR 200	1x Y6 and 1xY5	29 <sup>th</sup> June
Lincolnshire school*	NOR 128	1x Y5 and 1xY6	29 <sup>th</sup> June
School 53	NOR 99	1x Y5/Y6	3 <sup>rd</sup> July
School 82	NOR 336	1xY6, 1xY5/Y6 and 1x Y5	3 <sup>rd</sup> July
<b>Control schools</b>			
School 52 (Lincolnshire)			
School 54 (Lincolnshire)			
School 86 (Leeds)			

\*This Lincolnshire school withdrew from the intervention before the end of the project

## Appendix J: Train the Trainers Observation Schedule

### Observation Schedule: Training the Trainers

#### Observation

Developer's name:.....

Date: ...../...../.....

Observation	Notes
Was there a discussion about writing genres?	
Did the developer explain and model the use of the mnemonic IPEELL?	
Did the trainers have an opportunity to practise the use of the mnemonic IPEELL?	

Did the developer explain and model the use of graphic organiser or planning sheet?	
Did the trainers have an opportunity to practise the use of graphic organiser or planning sheet?	
Was a form of self-scoring explained?	
Did the trainer model positive self-talk?	

**Other Notes**



## Appendix K: Train the Teachers Observation Schedule

### Observation Schedule: Training the Teachers

Trainers .....

.....

Venue .....

Date .....

Observation	Notes
Was information provided about the background to the project and expectations of schools?	
Did the trainer talk about the importance of first hand memorable experiences?	
Did the trainer explain the links between memorable experiences and the writing process using IPEELL?	
Was there a discussion about writing genres?	
Did the trainer talk about the cold task?	

Was a model text used for analysis of purpose, structure and language features?	
Did the trainer explain and model the use of the mnemonics PAT and IPEELL?	
Did the trainer explain and model the use of the planning frame?	
Did the trainer demonstrate the development of a mark scheme and discuss self-scoring?	
Did the trainer model / discuss positive self-talk and motivational messages?	
Did the trainer explain graphing of scores?	
Did the trainer discuss goal setting?	
Was peer review and scoring of writing discussed?	

Did the trainer explain the next stages in the research project?	
--	--

Additional notes

## Appendix L: School Observation Schedule

### CEP / EEF Writing Project – Observation Schedule

Name of observer.....

Name of School ..... Class.....

Name of Teacher.....

#### Scoring System:

3 – Well covered / Good explanation / All aspects evident

2 – Partly covered / Satisfactory explanation / Most aspects evident

1 – Partly covered / Minimal explanation / Some evidence

0 – Not covered / No explanation / No evidence

Lesson Observation	Score	Notes
Was there reference to a <b>memorable experience</b> ?		
Was there discussion about the writing genre being covered in the session?		
Did the teacher model writing in the genre?		
Did the teacher explain / revise the mnemonics PAT and IPEELL?		
Did children use the mnemonics PAT and IPEELL?		

Did the teacher model using a planning frame?		
Did the children use a planning frame?		
Did the teacher model positive self-talk?		
Did the teacher encourage children to use positive self-talk?		
Did the teacher model scoring?		
<b>Aspects noted through evidence:</b> <i>Pupil workbooks, displays, pupil comments</i>		
Evidence of the mnemonics PAT and IPEELL?		
Evidence of the use of planning frames?		
Evidence of motivational messages / positive self-talk?		
Evidence of self-scoring?		

---

Evidence of the graphing of scores?		
Evidence of goal setting?		
Evidence of peer-scoring?		

**Overall evaluation:(please circle)**

**Good implementation**

**Satisfactory implementation**

**Unsatisfactory implementation**

## Appendix M: Teacher Interview Schedules

- What has facilitated the successful delivery of the intervention?
- What are some of the actual and potential barriers to its successful delivery? How can some of these be overcome?
- Did you attend one of the training days? Was the training helpful? In what ways could it be improved?
- In what ways have you made use of the IPEEL manual?
- What do you think are the major outcomes of the intervention on the pupils?
- Do some pupils gain more than others? Why?
- What are they benefitting mostly from? Why?
- What do you feel you are gaining out of the intervention?

## Appendix N: Pupil Focus Group Schedules

### Exploring thoughts and feelings about the 'writing project' intervention - pupils' enthusiasm and attitudes

Tell me about your IPEELL writing lessons.. (what do you do in writing? What do you write? etc.)

Do you enjoy lessons about writing? (what do you particularly enjoy?)

What do you find fun/helpful/hard? (in what ways? Why?)

What do you think about your writing? (are you struggling, do you feel you have improved, what has helped you etc.)

Can you tell me about PAT? (how do you use it?)

Can you tell me about IPEELL? (what does it mean and how do you use it?)

Can you tell me about marking your writing and graphing your results? (in what particular ways do you use marking and graphing? Has marking/graphing helped you in any way?)

Would you change anything to make it better/easier/more fun?

Have you taken any trips or visited any interesting places with your school lately? Have any interesting visitors come to your school recently? Did you enjoy those visits/experiences? What did you particularly enjoy?

Has your teacher asked you to write about these visits/experiences? Did you enjoy writing about your experiences? Was it useful? In what ways?



## Appendix O: Sensitivity Analyses of Primary Outcome Data

Table 16: Analysis of primary outcomes with Key Stage 1 writing scores as the only baseline variable.

Trials	Outcome	Raw means				Effect size		
		Intervention group		Control group		n in model (intervention; control)	Proportiona l Odds Ratio	Hedges g (95% CI)
		n (missin g)	Proportion (95% CI)	n (missing)	Proportion (95% CI)			
<b>Trial 1</b>	Writing (Expected Standard or higher)	1243 (106)	0.70 (0.68,0.73)	1222 (129)	0.73 (0.71,0.76)	2465 (1243,1222)	0.85 (0.53,1.35)	-0.09 (-0.35,0.16)
	Outcome	n (missin g)	Mean (95% CI)	n (missing)	Mean (95% CI)	n in model (intervention; control)	Hedges g (95% CI)	
<b>Trial 2</b>	Writing	1164 (289)	21.6 (21.4,22.0)	1032 (332)	209 (20.5,21.2)	2196 (1164,1032)	0.11(-0.13,0.34)	

Table 17: Frequencies and percentages of pupils according to the ordinal levels of key Stage 2 writing scores in trial 1

outcome	TRIAL 1		TRIAL 2	
	Intervention group	Control Group	Intervention group	Control Group
	% (n )	% (n/N )	% (n )	% (n )

<b>GDS</b>	9.1 (120 )	10.4 (135)	13.7(193)	12.7(169)
<b>EXS</b>	59.4 (788 )	60.7 (789)	56.6(798)	58.5(777)
<b>WTS</b>	26.5 (351)	23.2 (302)	24.3(349)	22.2(295)
<b>PKG</b>	3.9 (52)	4.7 (61)	4.2(59)	4.4(59)
<b>PKE</b>	0.9 (1)	0.3 (4)	0.3(4)	1.2(16)
<b>PKF</b>	0.2 (3)	0.6 (8)	0.6(8)	0.1(1)
<b>BLW</b>	0.1 (1)	0.0 (0)	0.0(0)	0.8(11)
<b>Total (N)</b>	1326	1300	1411	1328

**Table 18: Frequencies and percentages of FSM pupils according to the ordinal levels of key Stage 2 writing scores in trial 1**

outcome	TRIAL 1		TRIAL 2	
	Intervention	Control	Intervention	Control
	group	Group	group	Group
	% (n )	% (n/N )	% (n )	% (n )
<b>GDS</b>	4.4(22)	6.7(34)	9.0(48)	7.2(37)
<b>EXS</b>	54.6(275)	55.5(282)	53.4(285)	51.9(268)
<b>WTS</b>	34.7(175)	27.6(140)	29.0(154)	33.1(171)
<b>PKG</b>	4.6(23)	8.9(45)	7.1(38)	5.4(28)
<b>PKE</b>	1.6(8)	0.6(3)	0.2(1)	1.4(7)
<b>PKF</b>	0.2(1)	0.8(4)	1.3(7)	0.2(1)
<b>BLW</b>	0.0(0)	0.0(0)	0.0	0.8(4)
<b>Total (N)</b>	504	508	532	516

**Table 19: Frequencies and percentages of low achieving pupils according to the ordinal levels of key Stage 2 writing scores in Trial 1**

outcome	TRIAL 1		TRIAL 2	
	Intervention	Control	Intervention	Control
	group	Group	group	Group
	% (n )	% (n/N )	% (n )	% (n )
<b>GDS</b>	1.0(6)	2.1(11)	2.2(13)	1.0(5)
<b>EXS</b>	39.7(229)	4.9(223)	38.5(227)	37.0(195)

<b>WTS</b>	47.8(276)	42.7(227)	47.8(282)	46.1(243)
<b>PKG</b>	9.0(52)	10.9(58)	9.5(56)	10.6(56)
<b>PKE</b>	1.9(11)	0.8(4)	0.7(4)	3.0(16)
<b>PKF</b>	0.5(3)	1.5(8)	1.4(8)	0.2(1)
<b>BLW</b>	0.0(0)	0.2(1)	0.0(0)	2.1(11)
<b>Total (N)</b>	577	532	590	527

**Table 20: Frequencies and percentages of high achieving pupils according to the ordinal levels of key Stage 2 writing scores in Trial 1**

outcome	TRIAL 1		TRIAL 2	
	Intervention	Control	Intervention	Control
	group	Group	group	Group
	% (n )	% (n/N )	% (n )	% (n )
<b>GDS</b>	15.2(114)	16.1(124)	21.9(180)	20.5(164)
<b>EXS</b>	74.7(559)	73.7(566)	69.6(571)	72.7(582)
<b>WTS</b>	10.0(75)	9.8(75)	8.2(67)	6.5(52)
<b>PKG</b>	0.0(0)	0.4(3)	0.4(3)	0.4(3)
<b>PKE</b>	0.1(1)	0.0(0)	0.0(0)	0.0(0)
<b>PKF</b>	0.0(0)	0.0(0)	0.0(0)	0.0(0)
<b>BLW</b>	0.0(0)	0.0(0)	0.0(0)	0.0(0)
<b>Total (N)</b>	749	768	821	801

**Table 21: Regression coefficients of multilevel analysis of study specific writing scores in Trial 2**

Variables	Estimates	95% Lower bound	95% High bound
Intercept	2.69	0.53	4.85

KS1 Writing	1.26	1.19	1.32
Intervention	0.60	-0.56	1.75
LA (Leeds- Lincolnshire)	-0.68	-1.97	0.60
SchoolSize	0.00	0.00	0.01
Mixed Classes	0.25	-1.34	1.85
FSM	-0.73	-1.21	-0.26
<b>Covariance Parameters</b>			
School	5.83	-	-
Pupil	24.95	-	-
ICC	0.19	-	-

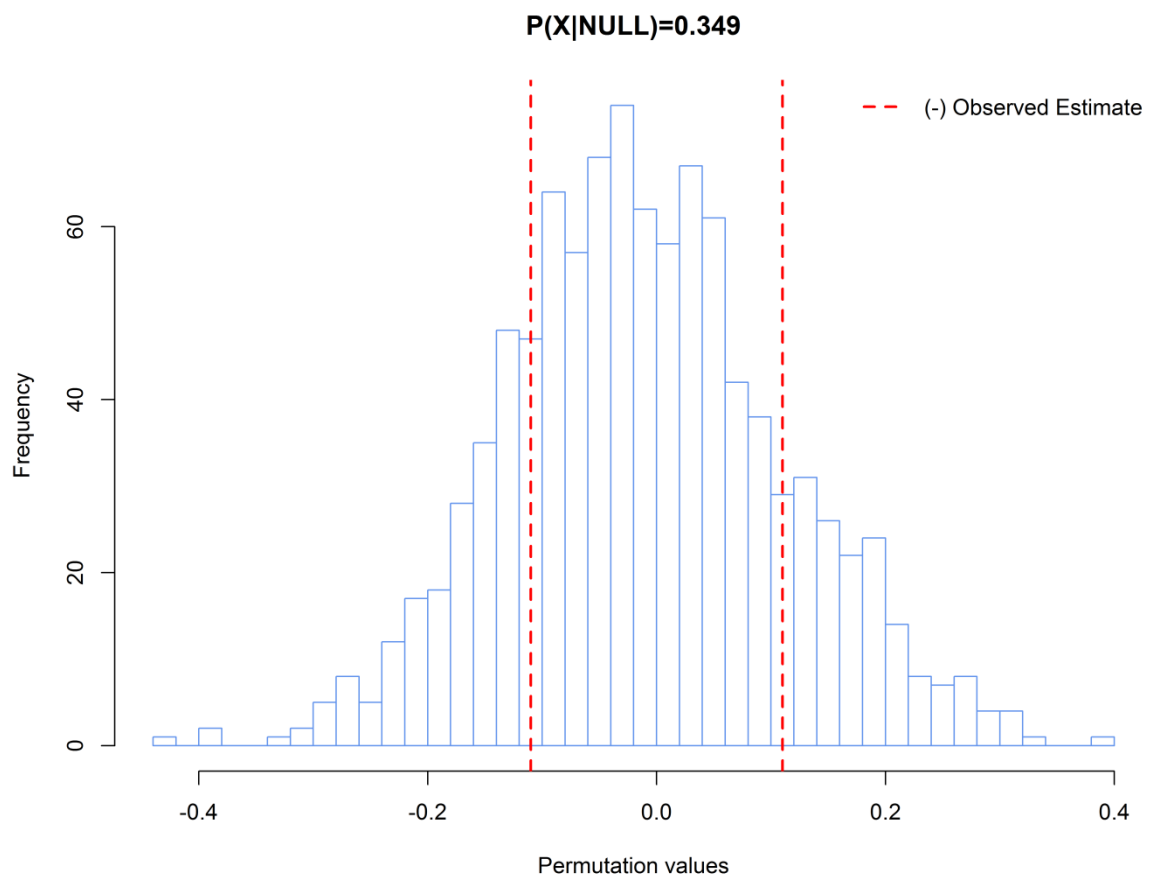


Figure 3: Distribution of effect size of primary outcome in Trial under the null hypothesis of no effect.

## Appendix P: Subgroup Analyses of Secondary Outcomes

Table 22: Subgroup analyses of secondary outcomes for low achievers defined by Key Stage 1 writing scores < 15.

Trial	Outcome	Raw means				Effect size		
		Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	
		n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
Trial 1	Maths	480 (25)	97.31 (96.69,97.93)	444 (22)	98.44 (97.79,99.08)	924 (480,444)	-0.24 (-0.53,0.05)	
	Reading	473 (32)	94.81 (94.13,95.48)	437 (29)	96.25 (95.55,96.96)	910 (473,437)	-0.28 (-0.55,-0.02)	
	Spelling (GPS)	482 (23)	96.82 (96.23,97.41)	450 (16)	97.66 (97.04)	932 (482,450)	-0.24 (-0.52,0.05)	
Trial 2	Maths	507 (29)	97.5 (96.8,98.2)	444 (36)	97.5 (96.7,98.2)	951 (507,444)	-0.29 (-0.56,-0.01)	
	Reading	508 (28)	96.3 (95.6,97.0)	439 (41)	95.8 (95.0,96.5)	947 (508,439)	-0.11 (-0.30,0.08)	
	Spelling	508 (28)	98.3 (97.7,98.9)	446 (34)	98.0 (97.4,98.6)	954 (508,446)	-0.23 (-0.49,0.03)	
	<b>Outcome</b>	<b>n (missing)</b>	<b>Proportion (95% CI)</b>	<b>n (missing)</b>	<b>Proportion (95% CI)</b>	<b>n in model (intervention; control)</b>	<b>Proportional Odds Ratio</b>	<b>Hedges g (95% CI)</b>
	Writing (Expected Standard or higher	531 (5)	39.4 (35.3,43.6)	470 (10)	37.7 (33.4,42.1)	1001 (531,470)	1.01 (0.63,1.63)	0.01 (-0.25,0.27)



Tri al	Outcom e	Raw means				Effect size	
		Intervention group		Control group		n in model (interventi on; control)	Hedges g (95% CI)
		n (missing )	Mean (95% CI)	n (missin g)	Mean (95% CI)		
Trial 1	Maths	746 (7)	104.51 (104.09,104.93)	771 (9)	105.11 (104.70,105.53)	1517 (746,771)	-0.18 (-0.45,0.09)
	Reading	747 (6)	104.68 (104.19,105.17)	772 (8)	105.11 (104.63,105.59)	1519 (747,772)	-0.19 (-0.38,-0.01)
	Spelling (GPS)	746 (7)	106.51 (106.09,106.93)	772 (8)	106.81 (106.40,107.22)	1518 (746,772)	-0.18 (-0.42,0.06)
Trial 2	Maths	819 (4)	105.1 (104.7,105.5)	782 (22)	106.3 (105.9,106.7)	1601 (819,782)	-0.29(-0.57,-0.02)
	Reading	820 (3)	105.9 (105.4,106.3)	782 (22)	106.6 (106.2,107.1)	1602 (820,782)	-0.17 (-0.36,0.01)
	Spelling	820 (3)	107.8 (107.4,108.2)	782 (22)	108.9 (108.5,109.3)	1602 (820,782)	-0.27 (-0.51,-0.03)
	Writing (Expecte d Standard or higher	821 (2)	91.5 (89.4,93.2)	801 (3)	93.1 (91.2,94.7)	1622 (821,801)	0.92 (0.61,1.38)

Table 23: Subgroup analyses of secondary outcomes for high achievers defined by Key Stage 1 writing scores  $\geq 15$

Table 24: Sensitivity analysis of primary outcome in Trial 1 using multilevel model

	Raw means				Effect size	
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)		
ALL pupils	1243 (106)	5.75 (5.71,5.79)	1222 (129)	5.78 (5.74,5.82)	2465 (1243,1222)	-0.09 (-0.33,0.15)
FSM	478 (38)	5.57 (5.50,5.64)	484 (44)	5.59 (5.52,5.66)	962 (478,484)	-0.06 (-0.35,0.24)
Low Achievers	494 (11)	5.29 (5.23,5.36)	454 (12)	5.32 (5.25,5.39)	948 (494,454)	-0.06 (-0.34,0.22)
High Achievers	749 (4)	6.05 (6.01,6.09)	768 (12)	6.06 (6.02,6.09)	1517 (753,780)	-0.06 (-0.29,0.17)

## Appendix Q: Missing Data Analysis for Primary Outcomes

As explained on page 37, Appendix Q shows that the effect size with and without multiple imputations are comparable. The effect size from multiple imputations of all pupils was 0.11 (-0.09, 0.27). The effect sizes for the pre-specified subgroup analyses were 0.12(-0.11, 0.32), 0.20(-0.01, 0.39) and 0.05(-0.17,0.25) for FSM, low achievers and high achievers, respectively.

**Table 25: The estimated odds ratios of missing data (writing outcome) in Trial 1 using logistic regression**

	All Pupils	FSM	Low Achievers	High Achievers
<b>Variables</b>				
Trial 1				
KS1 Writing	0.94 (0.87,1.02)	0.95 (0.87,1.04)	0.97 (0.85,1.11)	0.88 (0.65,1.19)
Intervention	0.63 (0.33,1.21)	0.55 (0.25,1.22)	0.87 (0.38,1.99)	0.36 (0.11,1.11)
LA (Leeds- Lincolnshire)	1.27 (0.64,2.53)	1.02 (0.46, 2.28)	1.12 (0.47, 2.68)	2.39 (0.76, 7.56)
SchoolSize	1.00 (0.99,1.01)	1.00 (0.99,1.01)	1.00 (0.99,1.01)	1.00 (0.99,1.01)
Mixed Classes	0.62 (0.27,1.44)	0.66 (0.24,1.82)	0.84 (0.27,2.64)	0.46 (0.13,1.67)
FSM	3.50 (1.67,7.32)	-	-	-

Table 26: The estimated odds ratio of missing data (writing outcome) in Trial 2 using logistic regression

Variables	All Pupils	FSM	Low Achievers	High Achievers
KS1 Writing	0.89 (0.86,0.94)	0.90 (0.86,0.94)	0.93 (0.88,0.98)	0.94 (0.87,1.01)
Intervention	0.68 (0.55,0.83)	0.72 (0.53,0.97)	0.51 (0.38,0.68)	0.89 (0.66,1.21)
LA (Leeds- Lincolnshire)	0.49 (0.39,0.61)	0.51 (0.37, 0.70)	0.53 (0.39,0.72)	0.55 (0.40, 0.77)
School Size	1.00 (0.99,1.01)	1.00 (0.99, 1.01)	1.00 (0.99,1.01)	1.00 (0.99,1.01)
Mixed Classes	1.27 (0.95,1.70)	1.54 (1.00,2.37)	1.26 (0.82,1.94)	1.41 (0.94,2.11)
FSM	1.55 (1.25,1.93)	-	-	-

Table 27: Sensitivity analysis for missing data using multiple imputations technique in Trial 2

Variables	ALL Pupils	FSM	Low achievers	High Achievers
Intercept	2.08 (0.24,3.93)	2.89 (0.54,5.24)	4.41 (2.18,6.64)	0.23 (-2.71,3.16)
KS1 Writing	1.25 (1.19,1.32)	1.19 (1.08,1.29)	1.01 (0.89,1.14)	1.37 (1.23,1.50)
Intervention	0.53 (-0.47,1.54)	0.67 (-0.58,1.92)	1.10 (-0.07,2.26)	0.30 (-0.88,1.48)
LA (Leeds-Lincolnshire)	0.65 (-0.43,1.74)	0.62 (-0.70,1.94)	0.74 (-0.42,1.90)	0.30 (-0.96,1.56)
School Size	0.00 (-0.00,0.01)	-0.00 (-0.01,0.00)	-0.00 (-0.01,0.00)	-0.00 (-0.01,0.00)
Mixed Classes	0.20 (-1.16,1.58)	-0.05 (-1.76,1.67)	-0.27 (-1.80,1.26)	0.17 (-1.43,1.78)
FSM	-0.53 (-1.00,-0.00)	-	-	-
Covariance Parameters				
School	4.09	4.30	3.26	5.19
Pupil	26.4	26.95	26.30	25.82
ICC	0.13	0.13	0.11	0.17

---

Effect Size	0.10 (-0.09,0.27)	0.12 (-0.11,0.32)	0.20 (-0.01,0.39)	0.05 (-0.17,0.25)
-------------	----------------------	----------------------	----------------------	----------------------

## Appendix R: Distribution of Pre-test Scores

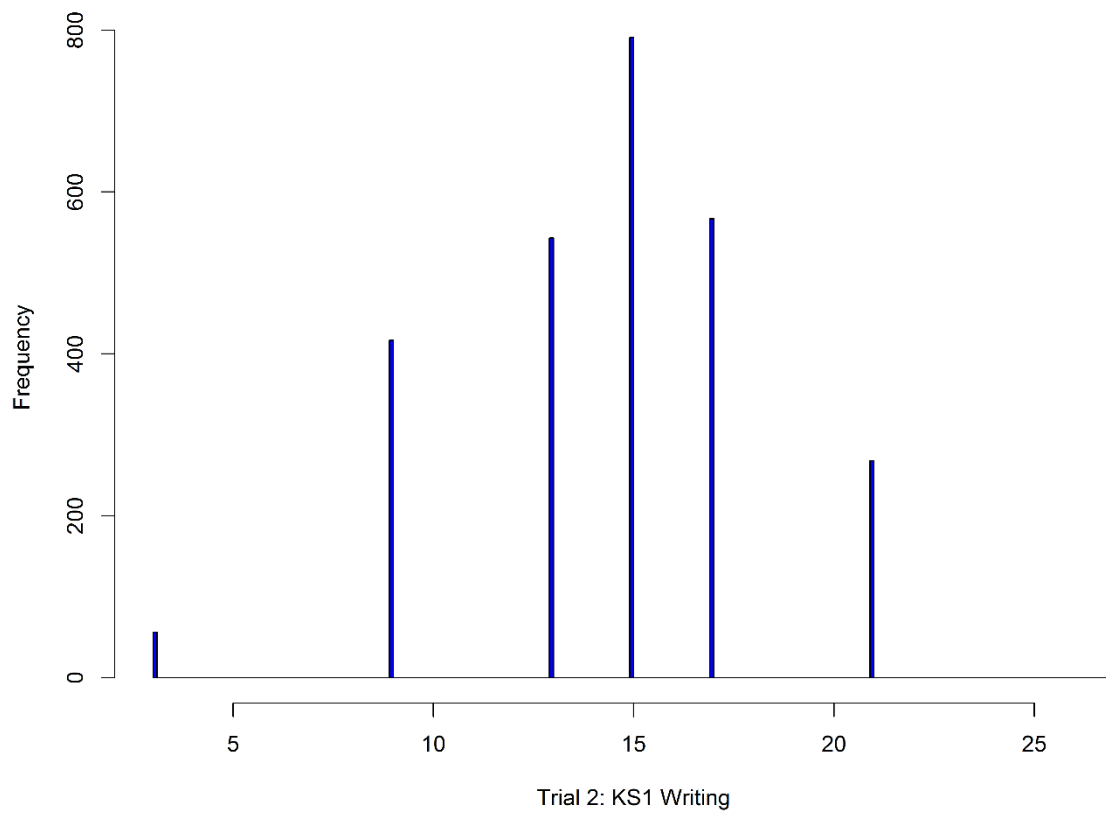
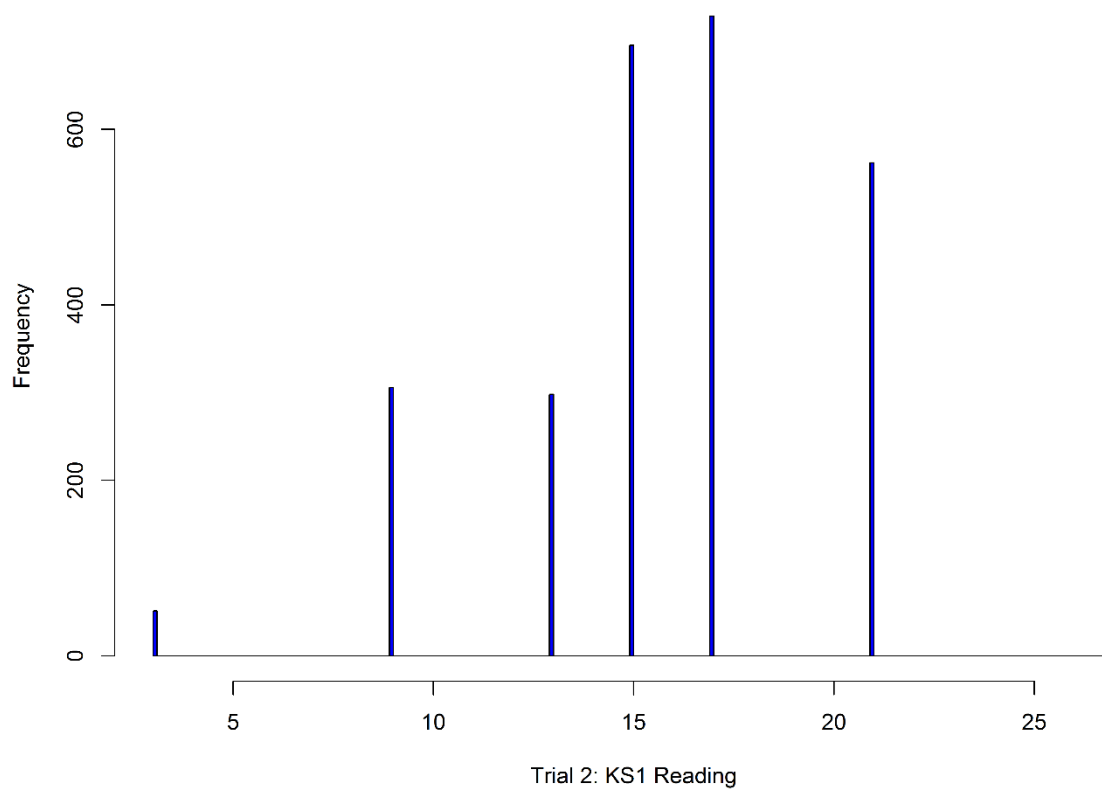


Figure 4: Distribution of KS1 writing in Trial 2



**Figure 5: Distribution of KS1 Reading in Trial 2**



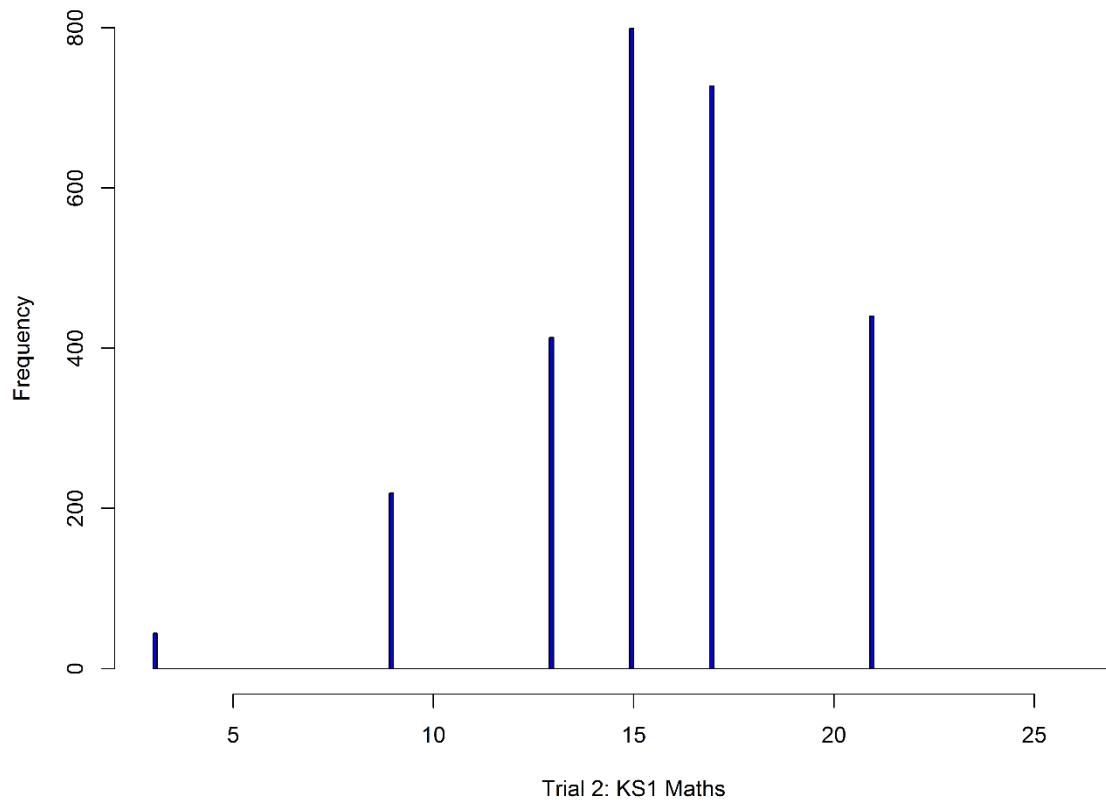
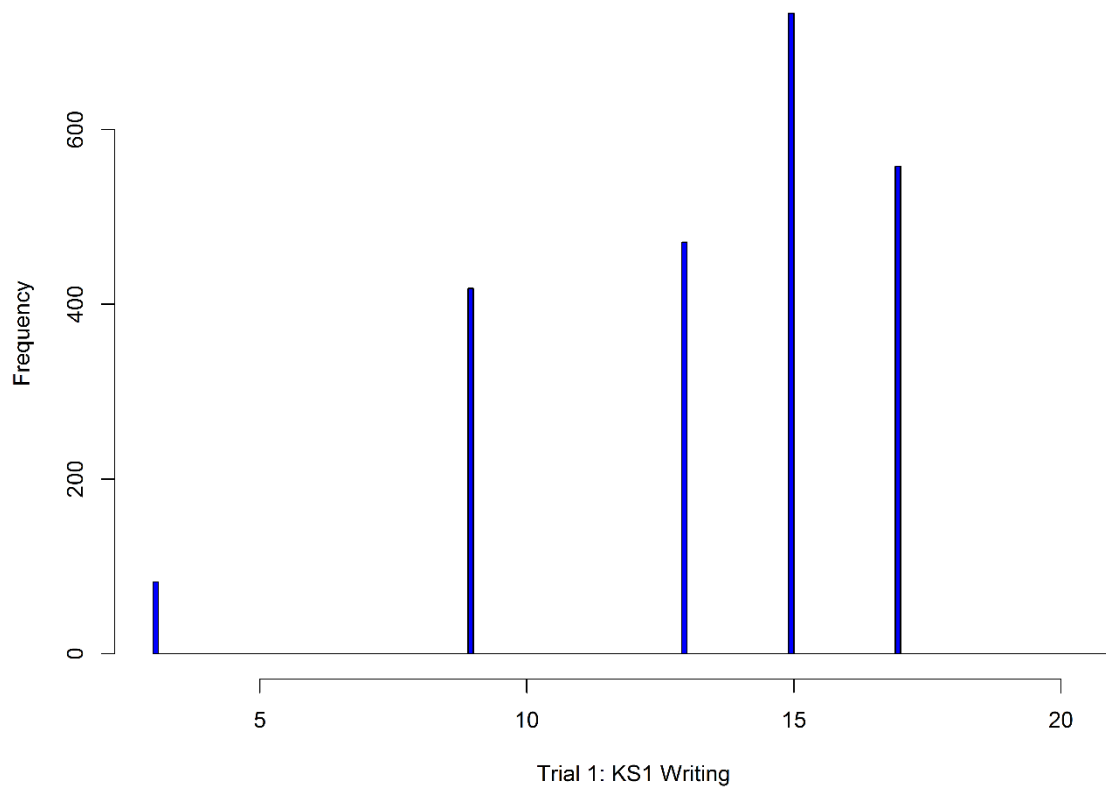


Figure 6: Distribution of KS1 Maths in Trial 2



**Figure 7: Distribution of KS1 writing in Trial 1**

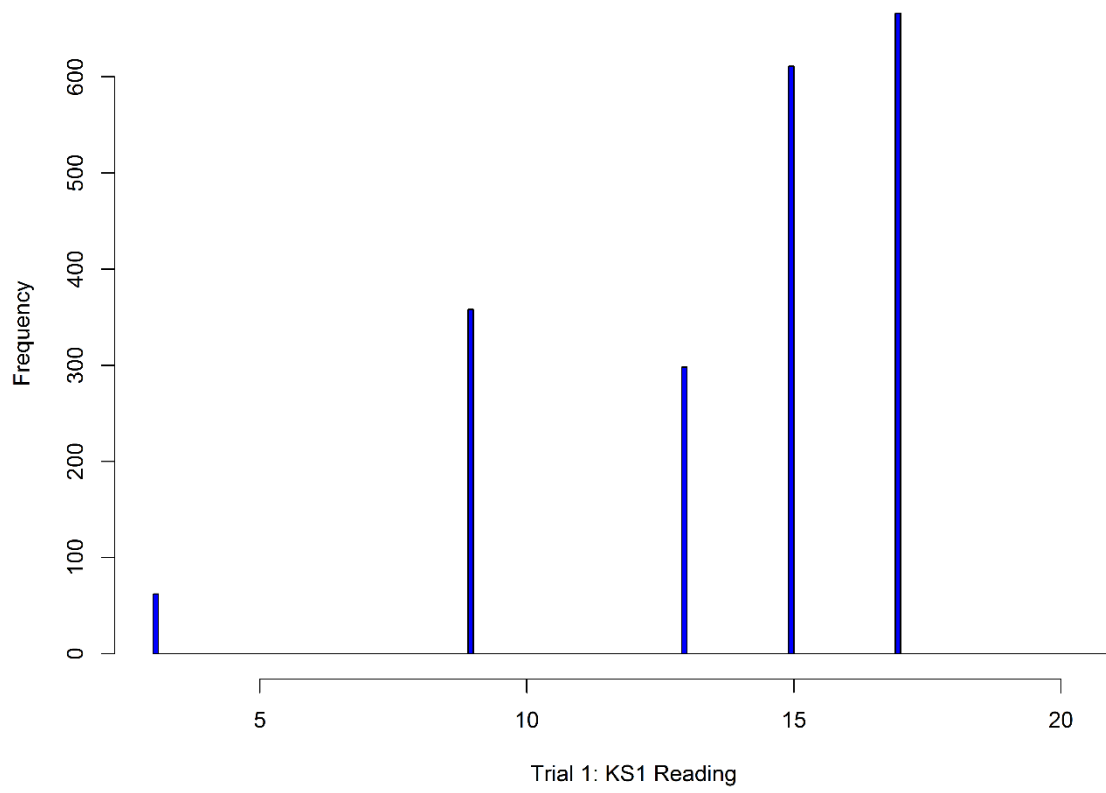


Figure 8: Distribution of KS1 Reading in Trial 1

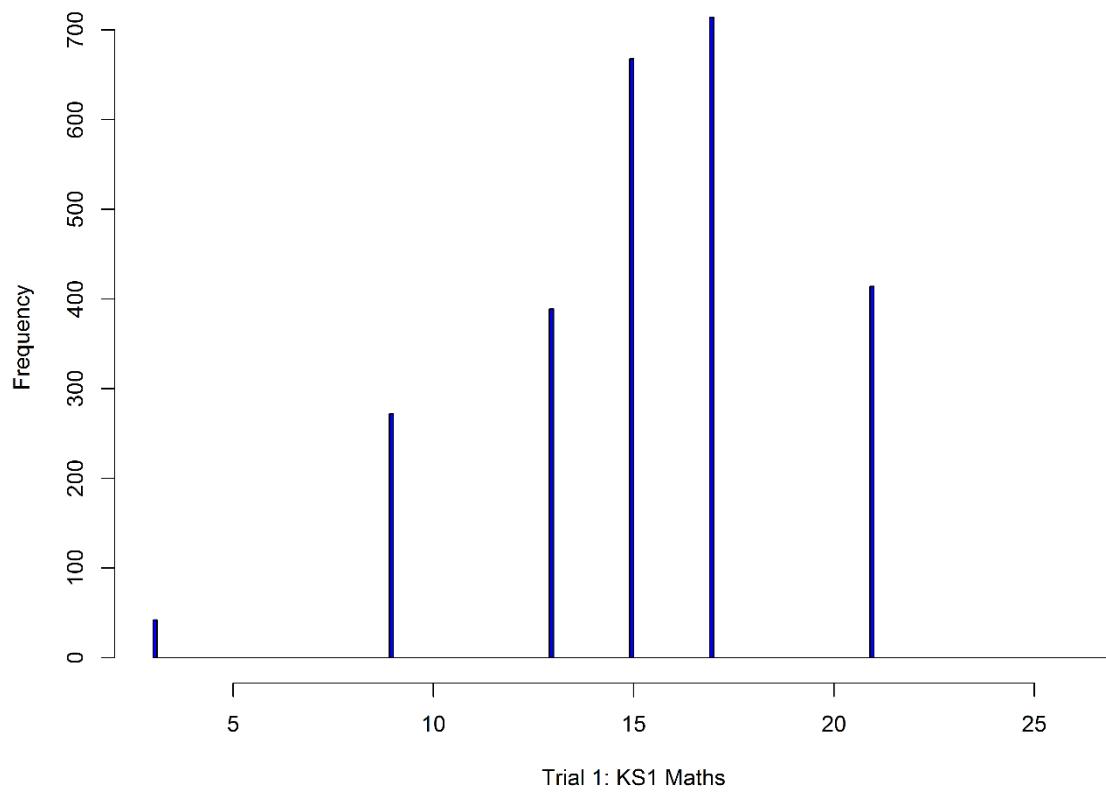


Figure 9: Distribution of KS maths in Trial 1

## Appendix S: Additional Results

Table 28: Multinomial analysis of the ordinal writing outcome in Trial 1, the reference group is a combination of PKG, PKEP, PKF and BLW because of the fewer cases in these categories

Parameters	GDS	EXS	WTS
	vs (PKG+PKEP+PKF+BLW)	vs (PKG+PKEP+PKF+BLW)	vs (PKG+PKEP+PKF+BLW)
<b>Odds Ratio</b>	0.87 (0.51,1.48)	0.97 (0.62,1.52)	1.22 (0.79,1.89)
<b>Effect Size</b>	-0.04 (-0.19,0.11)	-0.01 (-0.13,0.12)	0.06 (-0.07,0.18)

Table 29: Exploratory analyses of the impact of quality of implementation of IPEELL interventions, four categories of IPEELL intervention based on the fidelity scores were compared with the comparison group.

Trials	Quality of Implementation Ratings			
	Very Good	Good	Satisfactory	Unsatisfactory
<b>Trial 1</b>	-0.04 (-0.46,0.37)	-0.08 (-0.40,0.23)	-0.01 (-0.36,0.33)	-0.17 (-0.82,0.49)
<b>Trial 2</b>	0.30 (-0.08,0.69)	0.08 (-0.22,0.39)	0.03 (-0.28,0.34)	0.05 (-0.49,0.60)

Table 30: Characteristics of pre-test scores for the final analysis samples in Trial 1 and Trial 2

Trials	outcomes	Intervention group		Control group		Effect Size
		n (missing)	[Mean $\pm$ SD]	n (missing)	[Mean $\pm$ SD]	
<b>Trial 1</b>	KS1 Maths	1226	15.6	1215	15.4	0.03
		(123)	(3.6)	(100)	(3.6)	(-0.18,0.24)
	KS1 Reading	1220	15.6	1209	15.4	0.05
		(129)	(4.1)	(142)	(4.0)	(-0.16,0.25)
	KS1 Writing	1243	14.3	1222	14.3	0.0
		(106)	(3.9)	(129)	(4.0)	(-0.21,0.20)
<b>Trial 2</b>	KS1 Maths	1326	15.8	1226	15.6	0.08
		(127)	(3.3)	(138)	(3.6)	(-0.09,0.25)
	KS1 Reading	1328	16.0	1221	15.8	0.1
		(125)	(3.8)	(143)	(3.8)	(-0.07,0.20)
	KS1 Writing	1164	14.7	1032	14.8	0.0
		(289)	(3.6)	(332)	(3.6)	(-0.18,0.18)

## Appendix T: Distribution of Post-test scores

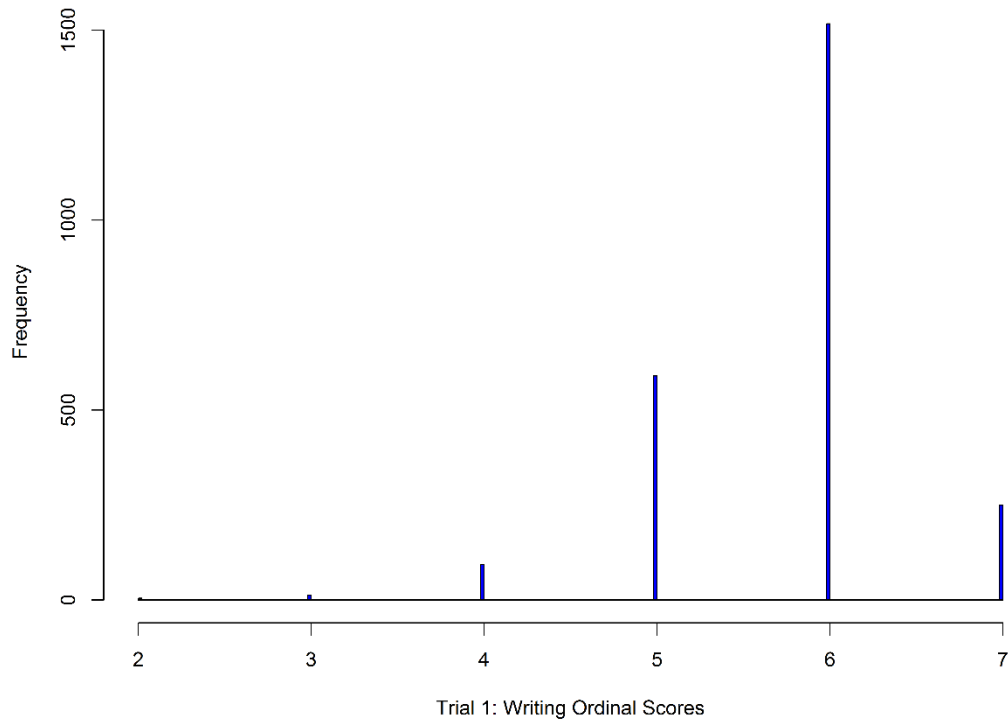


Figure 9: Distribution of Writing Scores in Trial 1

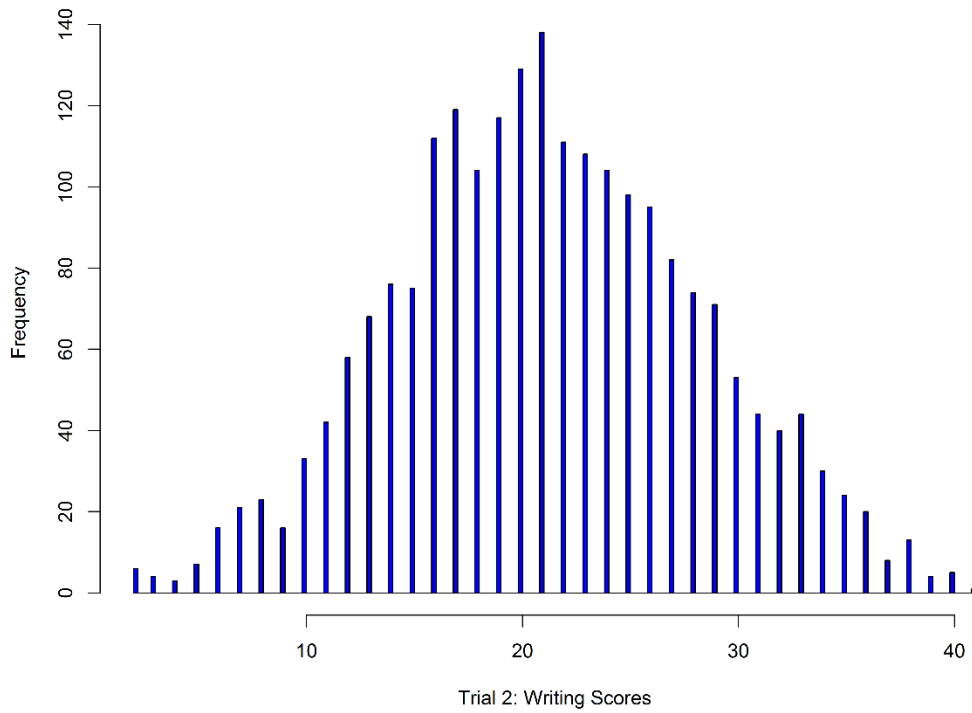
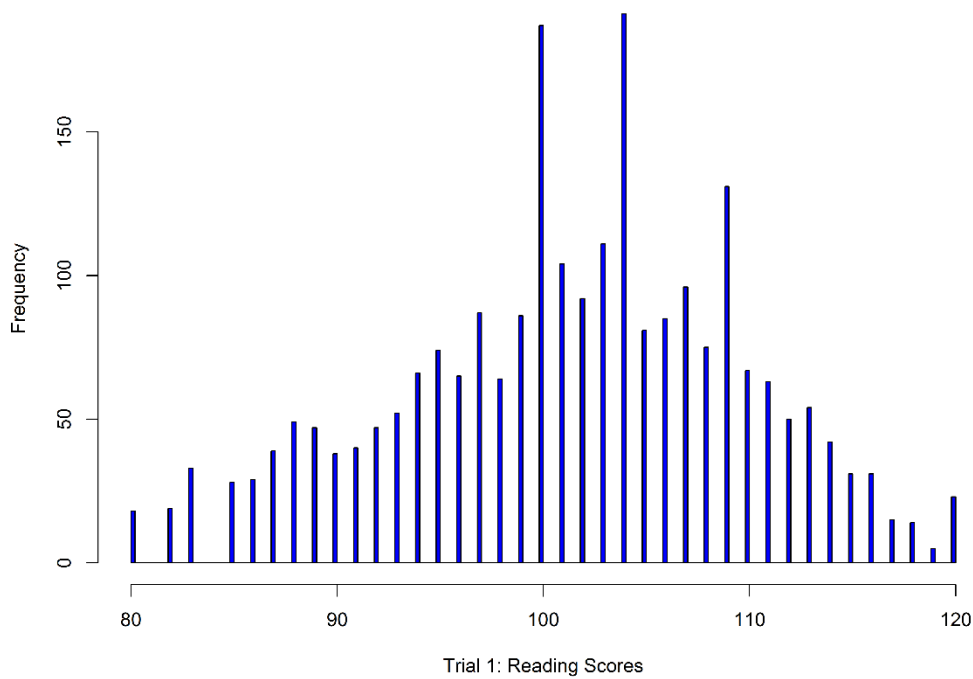


Figure 10: Distribution of Writing Scores in Trial 2





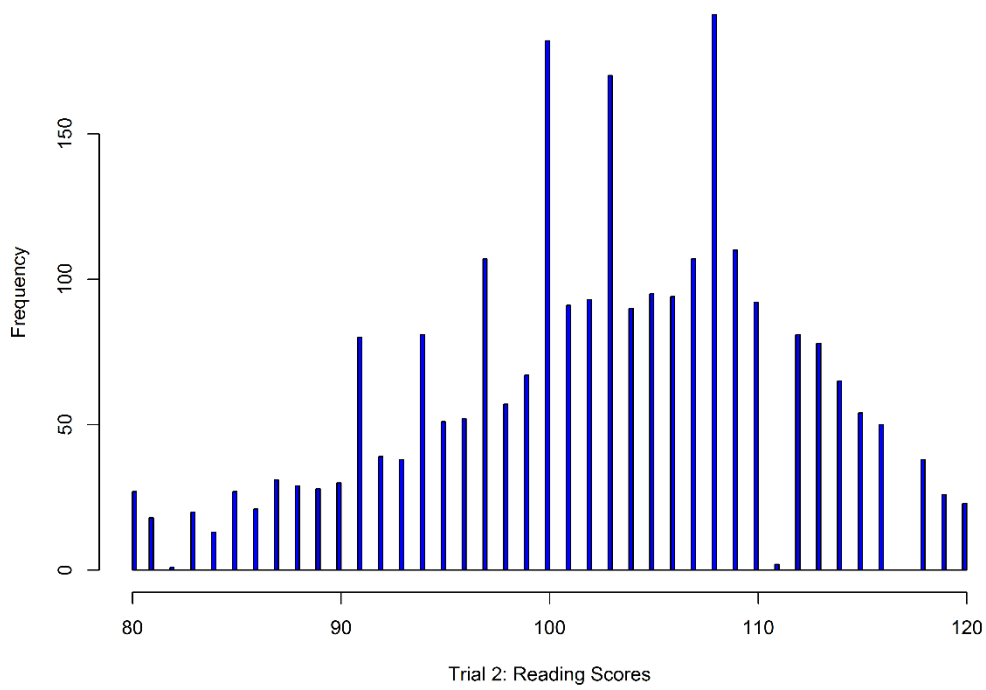


Figure 12: Distribution of Reading Scores in Trial 2

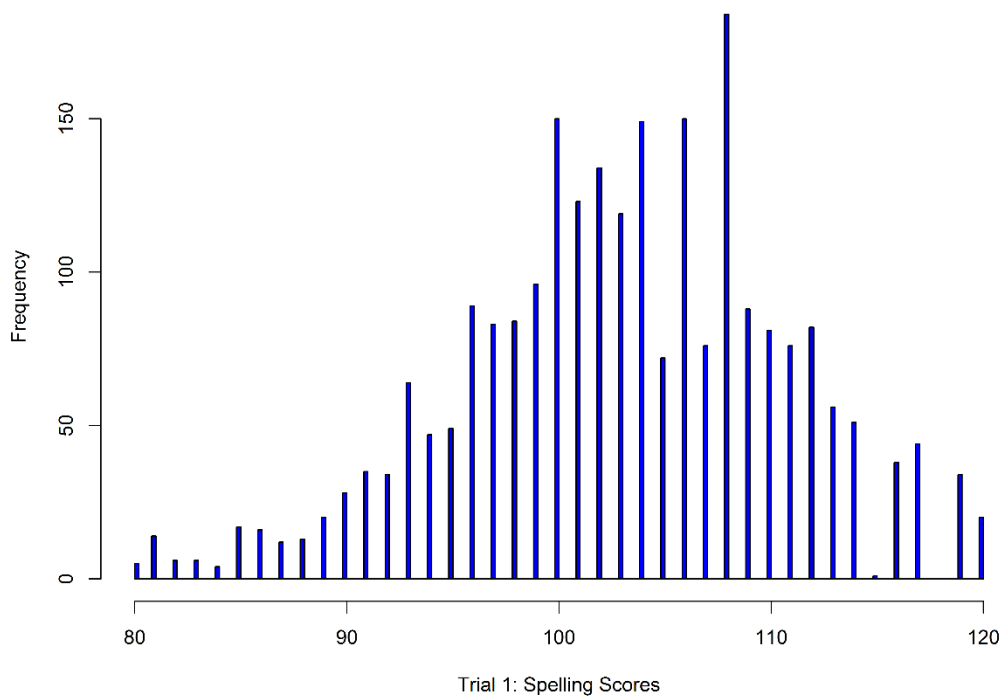


Figure 12: Distribution of Spelling Scores in Trial 1

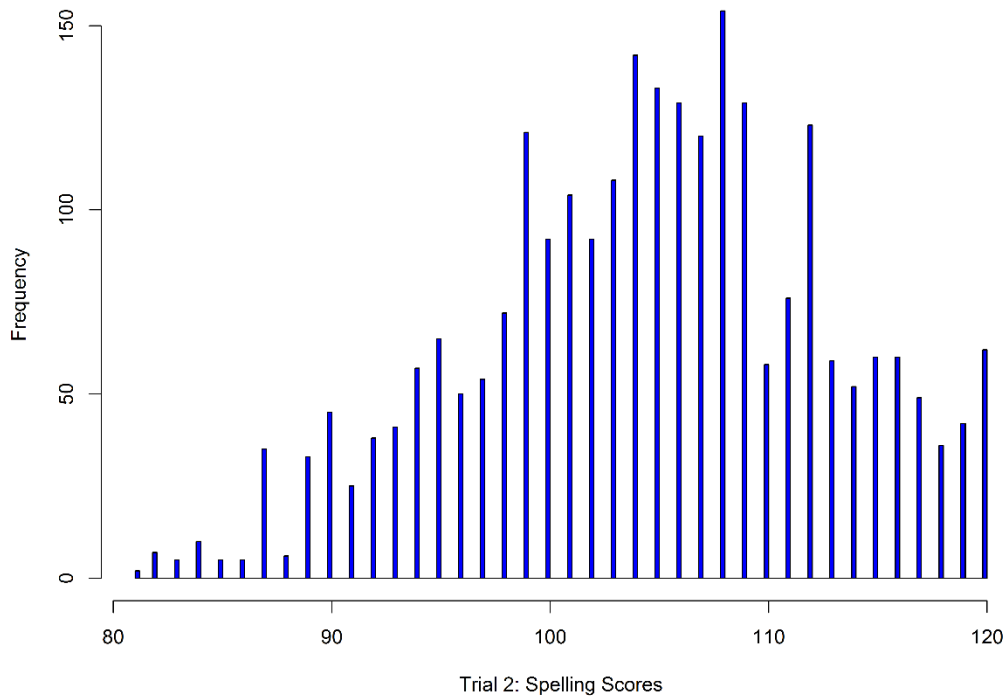


Figure 13: Distribution of Spelling Scores in Trial 2

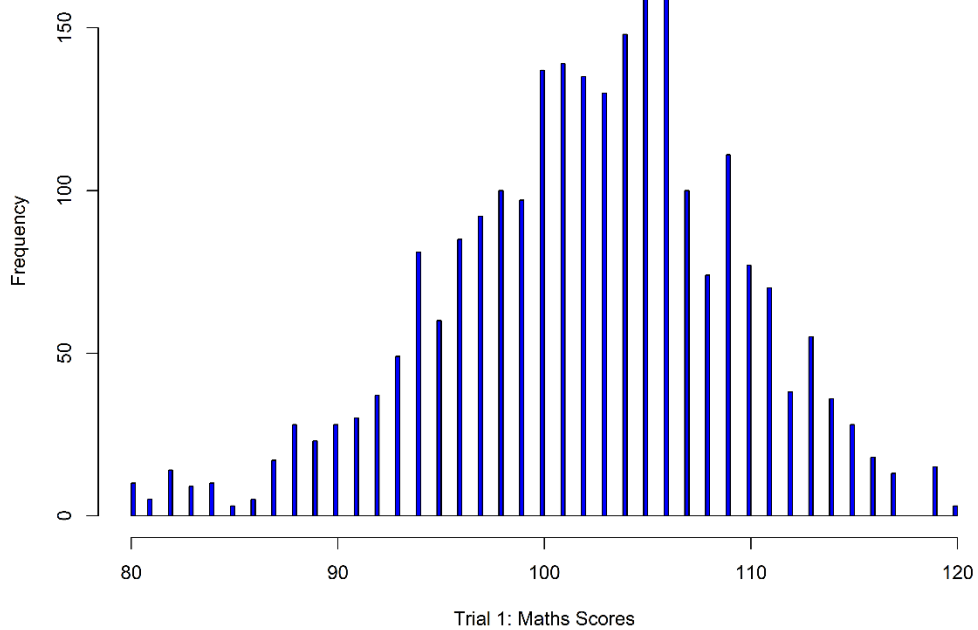


Figure 14: Distribution of Math Scores in Trial 1

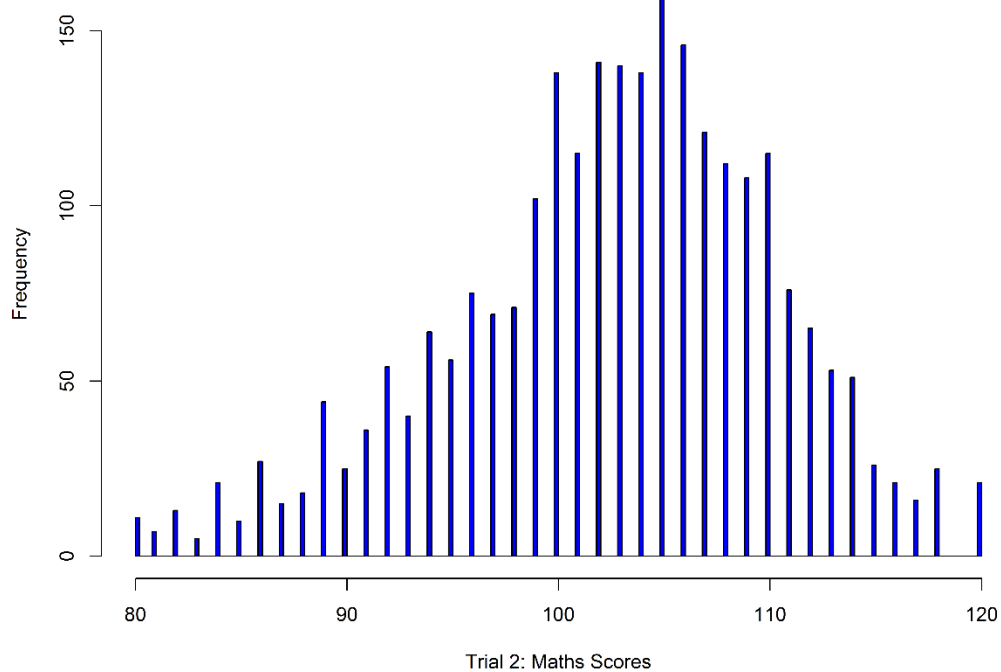


Figure 15: Distribution of Math Scores in Trial 2

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

**OGI** This information is licensed under the Open Government Licence v3.0. To view this licence, visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/>

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

This document is available for download at [www.educationendowmentfoundation.org.uk](http://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](http://www.educationendowmentfoundation.org.uk)