



Integrating English

Evaluation Report

September 2019

Martin Culliney, Nick Moore, Mike Coldwell and Sean Demack



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 (formerly Impetus Trust) 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.

The project was co-funded by The Bell Foundation and Unbound Philanthropy, as part of a funding round focusing on children with English as an additional language.

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

-  Jonathan Kay
Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP
-  0207 802 1653
-  jonathan.kay@eefoundation.org.uk
-  www.educationendowmentfoundation.org.uk



Contents

About the evaluator	4
Executive summary	5
Introduction	7
Methods	14
Impact evaluation.....	22
Implementation and process evaluation	31
Conclusion	51
References.....	54
Appendix A: EEF cost rating	57
Appendix B: Security classification of trial findings	58
Appendix C: Logic model	60
Appendix D: Distribution of baseline and outcome measures	61
Appendix E: Primary analyses and calculation of effect size	64
Appendix F: IPE survey.....	66
Appendix G: Observation schedule for training visits	71
Appendix H: Interview schedule for school visits.....	73
Appendix I: Memorandum of Understanding.....	76
Appendix J: Fair Processing Notice	80
Appendix K: Fidelity measures	85
Appendix I: Baseline survey responses.....	86
Appendix J: Pupil writing assessment (primary outcome)	87

About the evaluator

The evaluation team at Sheffield Institute of Education includes experts in trial design, analysis, implementation/process evaluation and linguistics/ESOL. The team has many years' experience of conducting evaluation studies as part of the Education Endowment Evaluation evaluator panel, and for organisations including government departments, charities, and other policymakers.

As well as the report authors, the project team included Helen Donaghue, Ester Ehiyazaryan-White, Diana Ridley, and Roberta Taylor who undertook fieldwork as part of the Implementation and Process Evaluation, and Linda Bray who was the project administrator. The report authors would also like to acknowledge the involvement of our team of test invigilators.

The evaluators would like to thank Sharon Davies at Enfield LA and Michelle Stanley, formerly at Enfield, for supporting and positively engaging with the evaluation throughout and to colleagues at the EEF and The Bell Foundation for advice and guidance. Finally, special thanks to the Integrating English delivery team, school leaders, teachers, and children who took part in the evaluation.

Executive summary

The project

The Integrating English programme aimed to improve the writing ability of Year 5 and 6 pupils through a structured CPD programme that provided teachers with linguistic and pedagogical strategies. All students were expected to benefit, but the greatest impact was expected on EAL pupils. The programme was centred on LiLAC, a training course that is widely used in Australia, which emphasises the importance of recognising differences between genres of text, using a 'teaching and learning cycle' to scaffold language, the 'register continuum' (thinking about the 'what', 'who', and 'how' of communication), and the relationship between spoken and written language.

The intervention spanned two years. In the first year, Year 5 teachers were trained for five days in the spring term before producing schemes of work which incorporated the training. These were delivered to pupils in the summer term of Year 5. During this summer term and the following autumn term, Year 6 teachers were then trained for five days before designing and delivering schemes of work to the same students, now in Year 6.

This was a randomised controlled trial involving 91 schools. Schools were randomly allocated to either receive the Integrating English programme, or to be in a 'business as usual' control. The process evaluation included observations of teacher training, case study visits to schools, and analyses of schemes of work. The trial took place between October 2016 and July 2018. The programme was developed and delivered by Enfield LA School Improvement Team, and was funded by the EEF, Unbound Philanthropy, and The Bell Foundation.

Key conclusions

1. There is no evidence that Integrating English improved pupils' KS2 writing outcomes. This result has a moderate to high security rating.
2. There is no evidence that Integrating English had an impact on the KS2 writing outcomes of pupils receiving free school meals. These results have lower security than the overall findings because of the smaller number of pupils.
3. There is no evidence that Integrating English improved EAL pupils' KS2 writing outcomes. Although this was measured through a large subgroup analysis, these results have lower security than the overall findings because of the smaller number of pupils.
4. The process evaluation indicates that, although teachers responded positively to the training, the CPD model may not have been effective in creating the desired teacher practice change. A simpler model may be more effective.

EEF security rating

This trial was an efficacy trial, which tested whether the intervention worked under developer-led conditions in a number of schools. It was a well-designed, two-armed randomised controlled trial, which was well powered. Twenty-four percent of the pupils who started the trial were not included in the final analysis because ten schools did not do the primary outcome test, and of the 80 schools remaining, 15% of their pupils did not sit the outcome test.

Additional findings

Children in the Integrating English schools made the equivalent of one months' less progress in KS2 writing, on average, compared to the control schools. The same finding was also identified for EAL pupils. In addition, the trial found that Integrating English had a small negative impact on KS2 Reading, and KS2 Grammar, Spelling and Punctuation. However, none of these findings were statistically significant.

Only a minority of schools completed the full demands of the programme (including training, creating action plans and schemes of work). Case study interviews as part of the process evaluation did indicate that training was positively received by the vast majority of interviewees; most interviewees felt that training was comprehensive and efficient and reported that they had shared strategies with other teachers. Interviewees also described how easily the aims of the training integrated with wider aims of the school, other initiatives and training programmes, as well as current practice.

However, although most interviewees mentioned the key programme principles when asked, the process evaluation indicated that key principles were addressed variably; some key processes and principles may not have been fully understood or delivered in the classroom. Exploratory analysis did not find additional benefit for pupils in schools that completed the full demands of the programme. The small numbers of pupils and non-random nature of the schools that were compliant mean that this finding should be treated with caution.


The process evaluation identified that schemes of work may be a useful mechanism for enacting professional development. While they were variably produced in this intervention, they may have the potential to act as a bridge between training and practice.

Cost

If delivered in geographical hubs, as in this trial, the average cost of the intervention is around £3 per pupil, per year when averaged over three years. Initial costs include training and buying the LiLAC manual, both of which are one off costs in the first year. Schools also need to meet the cost of staff cover for six days per participating teacher (12 days in total), for a Year 5 and Year 6 teacher. There are no ongoing costs associated with delivery. If an individual school wished to deliver the programme in a non-hub based model, due to higher training costs, the programme would cost around £21 per pupil, per year.

Impact

Summary of impact on primary outcome

Outcome/ Group	Effect size (95% confidence Interval)	Estimated months' progress	EEF security rating	No. of pupils	P value	EEF cost rating
KS2 Writing	-0.05 (-0.21; +0.12)	-1 month		3,607	0.577	£ £ £ £ £
KS2 Writing EAL pupils	-0.06 (-0.25; +0.12)	-1 month	N/A	1,978	0.517	£ £ £ £ £
KS2 Writing FSM pupils	+0.01 (-0.25; +0.12)	0 months	N/A	1,291	0.917	£ £ £ £ £

Introduction

Background evidence

Accurate and appropriate communication in English is recognised as a key performance indicator in education at primary and secondary level, and, of course, is the language of instruction in the English education system. Recognition of this has led to the introduction of policies creating targets for spelling and grammar tests at KS2 and GCSE English at KS4. Over the last ten years, debate on the role of grammar in the curriculum in U.K. schools has shifted from whether it should be taught to how it should be taught (Locke, 2010; Myhill and Watson, 2014). Students with English as an Additional Language (EAL) who demonstrate proficiency in English can close the attainment gap typically associated with students from minority backgrounds and exceed average attainment at school-leaving age (Strand, Malmberg and Hall, 2015; Demie and Strand, 2006). Similar discussions in Australia on minority students and grammar have prompted changes to the national curriculum, linking the teaching of grammar explicitly to social context and appropriateness (Derewianka, 2012). The resulting innovations in primary and secondary school grammar teaching have aimed to close the attainment gap for minority groups and significantly improve communication for all students (Rose and Martin, 2012; Macken-Horarik et al., 2015).

The intervention that is the subject of this report draws on curricular innovations in the Australian context. These innovations build on decades of pedagogic interventions and research in the use of language to improve literacy, particularly for disadvantaged students, combining the social realist sociology of Basil Bernstein (for example, 1990), a genre pedagogy approach to teaching and learning, and a systemic functional linguistic (SFL) perspective (Rose and Martin, 2012). From social realist sociology, the concept of 'recontextualisation' (Bernstein, 1975, 1981) is central to uncovering the strategies used by teachers and textbooks to 'translate' knowledge from disciplines into the primary and secondary classroom. Bernstein (1975) stressed the need to make pedagogies visible in order to reduce the advantages that students who share the dominant culture and linguistic habits of the school have over others. Genre pedagogy (Rose and Martin, 2012 and see below)—using the teaching-learning cycle of deconstruction–joint construction–independent construction—is one strategy that can make the recontextualisation of knowledge visible to all students in the class. Identifying patterns of variation of language with social context, or register, is a primary research goal of SFL (Halliday, 1978; Halliday and Matthiessen, 2004). General textual patterns (Martin, 1985), linguistic development in disciplines across primary and secondary subjects (Christie and Derewianka, 2008), and discipline-specific textual patterns and lexico-grammatical structures (Coffin, 2006; O'Halloran, 2004) have been identified as supporting genre pedagogy. It is these strands of research that have directly informed the LiLAC training programme.

The development of genre pedagogy in the Australian context has taken place since the late 1970s. Various reviews (Nebauer and Sungaila, 1980) have often focused on the responses of teachers to the programme, leading to the development of a 'good enough grammatics'—a toolbox that teachers can work with in the classroom without demanding a high level of linguistic knowledge (Macken-Horarik et al., 2018). However, evaluations of genre pedagogy, particularly the LiLAC (Language in Learning Across the Curriculum) programme (Custance, Dare and Polias, 2012) at the centre of the current evaluation, have thus far typically been case-based (Pomagalska, 2019), have focused on either writing or reading (Rose and Acevedo, 2006), or have used measures that limit reliable evaluation (Coffin, Acevedo and Lövestedt, 2013). The current project, Integrating English, is consequently the first large-scale (multi-site, multi-region) evaluation of a general implementation of genre pedagogy outside of Australia.

Integrating English was administered by the Enfield LA School Improvement Team following initial development with London Schools Excellence Fund (LSEF) support. The Enfield School Improvement Team ran a pilot study with ten schools from 2014 to 2015. In these schools, 24 teachers were trained in LiLAC, a training programme developed in South Australia independently of Enfield, LSEF and the evaluation team. LiLAC was used in the pilot project to support EAL learners in the classroom and in booster sessions by helping teachers recognise, discuss, and teach critical language features of different subjects. The pilot evaluation report was submitted in March 2015. It highlighted numerous examples of the project's success, including the range of experienced teachers taking part, the resources to support preparing materials and schemes of work, and ongoing guidance from mentors. The programme trialled was very different from this pilot in a number of ways. The pilot covered both primary and secondary schools and it involved extra tuition for EAL students in after-hours 'booster' lessons. To avoid unfair comparisons between extra tuition and none in an RCT, and to reduce costs for the large-scale trial, booster lessons were removed, and the current design

was negotiated to offer what was considered a fair chance for the programme to have an effect in a comparable context to control schools. Primary schools were chosen because they appear to provide greater comparability across contexts. This efficacy trial addresses the need for more robust evidence on how Integrating English works within the mainstream school system.

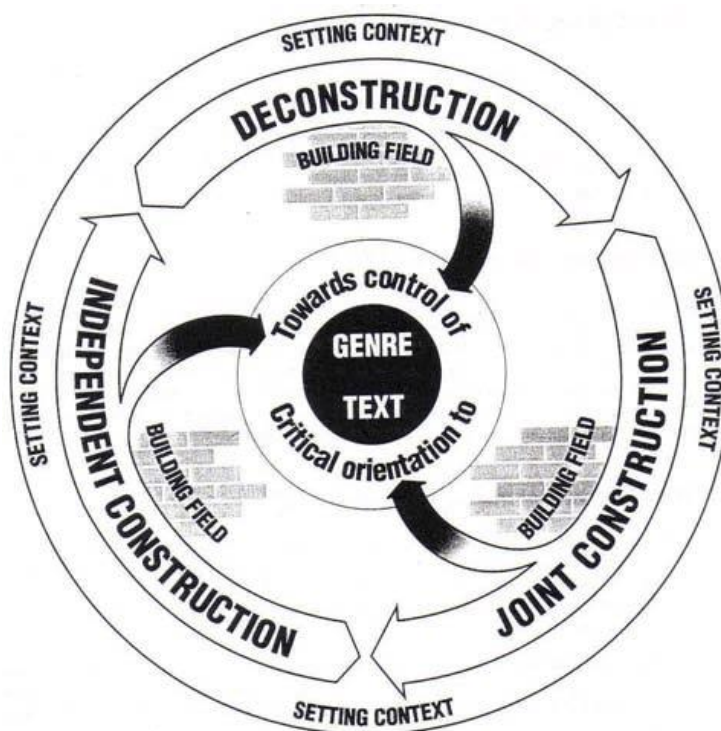
Intervention

The Integrating English programme is designed to improve the subject literacy and language pedagogy of teachers, and in turn improve the linguistic proficiency and communicative ability for all students. In this case, the programme was implemented with primary school teachers for the whole year group in treatment schools at Year 5 (summer term), in the first year of the intervention, and in Year 6 of the second year. The programme was expected to have greatest impact on EAL students, in turn leading to better learning across all school subjects for all pupils. The intervention is at school level so all children in the relevant year groups are taught using the Integrating English approach. All pupils were involved in testing except those whose parents indicated that they wished to withdraw.

The intervention is centred on LiLAC (Custance, Dare and Polias, 2012), a training course developed by Australian pedagogy and linguistics experts, widely used in Australia and owned by the Government of South Australia. The LiLAC course aims to enable teachers of any subject to adopt a functional approach to linguistics and grammar, aiming to break down the process of teaching the language of a subject to all pupils. It treats communication in academic subjects as a matter of ‘learning how to mean’ using the semiotic resources available to that subject; students learn through language, thus language learning is central to learning in all subjects (Halliday, 1993). A language has the potential to *mean* in all the social contexts that are (re)produced in a culture. One person’s use of language will reflect the cultural contexts where they are able to use linguistic resources to make meaning, whether this may be considered a first, second, or additional language, and so all learning is the expansion of a person’s potential to mean (Matthiessen, 2006). Central to this language-based approach to learning is the insight that language varies according to social context in conventional ways to create socially-recognisable ‘genres’ that achieve an established social purpose. Within a social semiotic theory of language, it is possible to identify these genres, including academic and pedagogic genres, and correlate the features of different social purposes with linguistic features (Martin, 1985). Social semiotic language programmes such as Integrating English encourage teachers to focus on language features and meanings that are made in each school subject. For example, teachers may focus on the generic and grammatical features of verbal art and everyday rhetoric (English); classification, experimentation, and reporting (science); recounts and causation (history); or problems, explanations, and proofs (mathematics), and so on (Coffin, 2006; O’Halloran, 2004).

The LiLAC programme also has a specific pedagogical approach designed to place texts at the centre of teaching and learning (Rose and Martin, 2012). The teaching–learning cycle of Setting Context, Deconstruction, Joint Reconstruction, and Individual Reconstruction (see Figure 1) offers opportunities to scaffold the learning of linguistic features relevant to a genre, and so is relevant to the different literacy demands of the various subjects in the school curriculum (Macken-Horarik et al, 2015). The pedagogical steps investigate the social context of the spoken, written, or multimodal text (Setting Context), break the text down to highlight genre-specific language features (Deconstruction), provide opportunities to investigate that language collaboratively (Joint Construction), and then support students as they apply that language to a modified context (Independent Construction).

Figure 1: The teaching–learning cycle in Genre Pedagogy (Rose and Martin, 2012)



Key to determining the literacy demands of school subjects is the register expected from students—or the variation in the Field (the ‘what’), the Tenor (the ‘who’), and the Mode (the ‘how’) of any communication (Halliday and Hasan, 1989). To respond appropriately within a social context, it is useful for students to think about where on a continuum the language of a text should be positioned—between the technical or abstract and the everyday or concrete; between the informal or personal and the formal or impersonal; and between the spoken or here-and-now and the written or decontextualized—to produce the text’s Field, Tenor and Mode, respectively (Custance, Dare and Polias, 2012). One pedagogic outcome of this perspective on language is the critical variations in language between spoken and written modes which enable the vernacular of pupils in any classroom to be valued in its own right but to be available for comparison with formal varieties. In suitably-designed pedagogic tasks, this distinction encourages students to speak in order to write (talk-as-process) or write in order to speak (talk-as-performance). As a consequence of this rich theory of the role of language as a social semiotic, the Integrating English programme highlights four key principles that successful participants are expected to understand and implement: the varieties of genre in texts in school subjects, the teaching-learning cycle, the register continuum, and structured talk (‘talk-as-performance’ or ‘talk-as-process’).

In addition, the Integrating English programme provides strategies for understanding and responding to the needs of EAL students in mainstream classes in the U.K. context, offering teachers a range of ready-to-use tips and techniques. During the intervention, teachers and trainers regularly referred to LiLAC techniques and strategies, whether they were part of the original LiLAC programme or variations introduced by the Integrating English programme. Through explorations of what it feels like to be surrounded by an unfamiliar language and discussions around the use of the child’s first language in the mainstream classroom, teachers were made aware of different responses to EAL students in class. Classroom activities that are introduced and often experienced include dictogloss, where a text is dictated at normal reading speed enough times for students to write down key words and ideas, but not enough to replicate the original text (Prince, 2013). Students then build these notes into a complete text of their own before finally comparing to the original to notice any differences in structure.

The Integrating English course was delivered by qualified teacher trainers recruited by the Enfield programme team. Five one-day accredited modules were attended by teachers in person, with readings and practical homework tasks between each session. Training was delivered through five regional ‘hubs’ using non-school venues, with teachers visiting their nearest regional centre for the training. The five hub areas were (1) South East, (2) West Midlands, (3)

North East, (4) Leicester and Peterborough, and (5) Cambridgeshire and Norfolk. The hub areas were not decided in advance but arranged once recruitment was complete to account for the location of schools signed up to the trial.

LiLAC training is an integral part of the Integrating English programme, but it is adapted to the U.K. context and supplemented by ongoing contact and support from trainers via email and FRONTER,¹ the online virtual learning environment (VLE) chosen by the delivery team specifically for this project (following successful implementation during the pilot programme), which is also used as a file-sharing resource for all teachers and trainers on the project. In addition, during the latter stages of training (on day five of the training), schemes of work (SoWs) are developed by trainees and trainers, based on learning from the Integrating English programme. Teachers are required by the programme to develop further SoWs back at their schools. Teaching based on these schemes of work then takes place in normal classroom time.

The intervention spanned two school years, as detailed below in Table 1. Phases one to three were aimed at Y5 teachers and classes. Phases four to six repeated these phases but were aimed at Y6 teachers and classes.

Phase 1 started for cohort 1 in January 2017 and involved two or more teachers (in Y5 in the first instance) from each school undertaking Integrating English training. The entire Integrating English course including training, readings, and tasks equated to approximately 50 hours of professional development for the teachers over the school term. Phase 2 followed the training, and again took place over half a term. Y5 teachers produced schemes of work informed by the Integrating English course, again supported by trainers. Phase 3 began, for cohort 1, after Easter. During the summer term, Y5 pupils in the two classes taught by the participating teachers were taught using methods based on the scheme of work developed in the previous phase.

Phase 4, starting while Phase 3 was running, comprised Integrating English training for the teachers with Y6 classes (the same cohort of pupils, but in the next academic year) in 2017/2018. This started in June 2017 and was completed at the start of the 2017/2018 Autumn term (see Table 1, below). Phase 5 saw schemes of work for Y6 classes developed over the remainder of the first half term of the 2017/2018 academic year. In Phase 6, Y6 pupils received teaching based on the Integrating English training and subsequent scheme of work for the Spring term 2017/2018. While teachers were not told to avoid using ideas from Integrating English in the Summer term 2017/2018, they were not required to plan at least 3 'LiLAC' lessons per week and it was assumed that that term would largely be devoted to preparing pupils for standardised tests. The teaching differs from normal practice in that the scheme of work and related teaching is shaped by the genre pedagogy approach underpinning this intervention.

In both Y5 and Y6, the intervention took place over two school terms. In term 1, teachers took Integrating English training and developed schemes of work. In term 2, teaching based on these schemes of work was delivered, although it is likely that some lessons using principles, techniques, or knowledge from Integrating English took place outside of term 2. The phasing of this programme across Y5 and Y6 as described above was adapted to fit the school calendar with the project timeline. This schedule also allowed for more teaching to be delivered using the schemes of work during the intervention period.

Table 1: Schedule of intervention delivery

	Y5 teachers	Y6 teachers	Pupils (2016/2017 Y5 cohort, Y6 in 2017/2018)
Spring term 2016/2017	1. receive LiLAC training 2. write scheme of work		
Summer term 2016/2017	3. implement scheme of work	4. receive some LiLAC training	teaching experienced by pupils in Y5
Autumn term 2017/2018		4. (continued) receive LiLAC training 5. write scheme of work	
Spring term 2017/2018		6. implement scheme of work	teaching experienced by pupils now in Y6

¹ For more information please see: <http://www.itslearning.eu/about-fronter>

In total, pupils in the treatment schools experienced lessons (intended to be three or more lessons per week) designed on Integrating English principles for at least two terms—one in the summer term of Y5 and one in the spring term of Y6—with different teachers in nearly all cases and often in different subjects from Y5 to Y6. For example, a pupil may have experienced Integrating English lessons in science for one term with their Year 5 teacher and Integrating English lessons in history for one term with their Year 6 teacher. The programme logic model (Appendix 3), developed with the delivery team, lays out the implementation logic of the programme and the causal mechanisms behind it. There were two linked processes:

- the implementation of a structured training programme was designed to change teacher practices in the classroom via the development of a new scheme of work (causal process 1); and
- the changes in teacher practices were expected to lead to improved pupil literacy outcomes and an embedded change in school practices (causal process 2).

KS2 Writing was the original choice as primary outcome. However, changes to the marking scheme in 2016 meant that results would only be available in categorical format, with pupils graded at one of three levels. This rendered it unsuitable for this evaluation, which required a scale variable as a primary outcome. The decision was therefore taken to administer an additional writing test in the summer of 2018 to assess the impact of the intervention. However, the test was only completed in 80 of the 90 schools participating in the trial. Full details on the outcome measures are supplied later in this report.

The original recruitment target of 100 schools was not met, and the trial commenced with 91 schools including control and treatment groups (one school withdrew from the project immediately after randomisation, leaving 90). Key characteristics of the trial include the levels of withdrawal and attrition of schools and the low rate of full compliance with the programme team's requirements for completion of the treatment. These are discussed further below.

Evaluation objectives

This evaluation was led by the Sheffield Institute for Education and funded by the EEF, the Bell Foundation, and Unbound Philanthropy. A clustered randomised controlled trial design (CRCT) was adopted to determine whether the intervention improves attainment in schools, building on the findings of the pilot study into this intervention, using a robust CRCT approach to strengthen the existing evidence base. Another important part of the evaluation was to verify that the programme was properly implemented by teachers with sufficient understanding of the model of language. The impact evaluation assessed the effect of Integrating English on the writing ability of all pupils, with subgroup analysis on EAL pupils and pupils entitled to free school meals (FSM). Writing ability was measured through a writing test administered by the evaluators using KS2 Writing past papers. The process evaluation sought to establish how the Integrating English programme was implemented in schools and to elicit the views of teachers on the efficacy of the intervention.

The Impact Evaluation addressed the following questions:

1. What is the impact of the intervention on the language ability of Y6 pupils, as measured by a writing test (KS2 past paper) as primary outcome, with KS2 Reading and KS2 English Grammar, Punctuation and Spelling (GPS) as secondary outcomes.
2. What is the impact of the intervention on EAL students, as measured by a writing test (KS2 past paper), with KS2 Reading and KS2 GPS as secondary outcomes.
3. What is the impact of the intervention on FSM students, as measured by a writing test (KS2 past paper), with KS2 Reading and KS2 GPS as secondary outcomes.
4. Where variations in implementation (from provision of training to delivery and final outcomes) occur, what effect does this have on the impact of the intervention?

The research questions for the Implementation and Process Evaluation (IPE) were:

1. How effective is the Integrating English programme in developing teachers' knowledge and understanding of language, based on systemic functional linguistics, in different subjects?
2. What evidence is there that this knowledge and understanding leads to improved classroom practice?
3. What evidence is there that this knowledge and understanding leads to improved pupils' language performance in the classroom?
4. What issues of fidelity occur during the trial?
5. What does the trial indicate about scalability?
6. What are the intervention costs?

Protocol:

https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/Integrating_English_Protocol_Amended_2017.11.07.pdf

SAP:

https://educationendowmentfoundation.org.uk/public/files/Projects/Integrating_English_SAP_2017.11.07___FINAL.pdf

Ethics and trial registration

The study design was approved by the SHU ethics committee in 2016.

Recruitment for this trial was overseen by the developers, Enfield LA School Improvement Team, and took place at the school level. The developers were responsible for promoting the project and discussing it with schools during the recruitment period. Schools signalled their commitment to take part in the trial by returning signed copies of the Memorandum of Understanding (Appendix I) to the Enfield team who compiled a master spreadsheet with the details of each participating school. The file was updated and shared with the evaluators at regular intervals. The next step in the recruitment process was the collection of pupil data. Again this task was undertaken by the developers, who received the data directly from schools and forwarded the individual files to the evaluators for merging into a single dataset. In instances where the data was incomplete or contained errors, both the evaluators and developers liaised with schools in order to ensure that the necessary data was obtained. Cooperating to share the workload in this manner was crucial for successfully concluding the recruitment phase ahead of randomisation.

The trial is publicly registered under International Standard Randomised Controlled Trial Number (ISRCTN) ISRCTN44415239: <http://www.isrctn.com/ISRCTN44415239>

Data protection

The project began in 2016, prior to the introduction of GDPR. Schools were informed about the purpose of data processing through participant information sheets and MoU (Appendix I). Both were revised later in the project and again sent to schools. The primary outcome assessment, which was administered in schools during June and July 2018, was coordinated by the evaluation team and marked by an external contractor. A data sharing agreement was signed to cover the temporary role of the contractor as data processor. Schools were informed of this arrangement through a Fair Processing Notice which was distributed before any data was shared or any assessments undertaken by pupils.

In the Fair Processing Notice it was stated that personal data would be retained for research and knowledge exchange purposes for a period of ten years after the last publication arising from the evaluation, with the longer-term archival value of the data to then be reviewed. Specifically, the personal

data shared between the evaluation team and the organisation contracted to provide marking services consisted of pupil names and their schools. No other identifying information was shared, and no special category data was shared. The legal basis for processing personal data was research as a 'Public Task' in accordance with GDPR Article 6 (1e).

Project team

Evaluation team (all Sheffield Hallam University)

- **Lead Director (Principal Investigator):** Professor Mike Coldwell, *Director of the Centre for Development and Research in Education*
- **Project manager and quantitative data analyst (Co-investigator):** Dr Martin Culliney, *Research Fellow*
- **Systemic functional linguistic expert and process evaluation lead (Co-investigator):** Dr Nick Moore, *Senior Lecturer*
- **Statistical consultant (Co-investigator):** Sean Demack, *Principal Research Fellow*
- **Process evaluation fieldworkers:** Dr Helen Donaghue, *Senior Lecturer*, Dr Diana Ridley, *Senior Lecturer*, Dr Roberta Taylor, *Senior Lecturer*, Dr Ester Ehiyazaryan-White, *Lecturer*

Delivery team

- **Project manager:** Sharon Davies, *Community Learning and Schools Programme Leader*, Skills for Work Service, Enfield Council
- **Project lead and LiLAC tutor:** Michelle Stanley, *Senior Teaching and Learning Consultant*, *School Improvement Service, Education Services – Schools and Children's Services*, Enfield Council

Methods

Trial design

The impact evaluation design was a two-arm, school-level clustered randomised controlled trial (CRCT). School level randomisation was used as it is more practical for recruitment and implementation. It also eliminates the risk of overspill, which is crucial for ensuring robustness. The control condition was 'business as usual'. Control schools were given £200 for taking part.

The original plan to use KS2 Writing as the primary outcome was abandoned in Autumn 2016 due to changes in the way it is recorded in the National Pupil Database, as described above. It was decided between SHU, Enfield, and the EEF that a bespoke writing assessment should be conducted, with the scores collected in scale format. The agreement was to use a KS2 Writing past paper and use pupil scores on this assessment as the primary outcome measure. At this stage, it was also agreed that KS2 Reading would be used as an additional secondary outcome. This report presents results from the primary outcome and both secondary outcomes for each analysis model.

Table 2: Trial design summary

Trial type and number of arms		Two-arm, school-level clustered randomised controlled trial (CRCT)
Unit of randomisation		Schools
Stratification variable(s) (if applicable)		Mean KS2 attainment, % EAL pupils, mean EAL fluency scale score, number of Y5 classes, geographical hub
Primary outcome	variable	Writing
	measure (instrument, scale)	Raw scores on past KS2 Writing papers (0–31 scale)
Secondary outcome(s)	variable(s)	KS2 Reading; KS2 Grammar, Punctuation and Spelling (GPS)
	measure(s) (instrument, scale)	NPD field: KS2_READMARK, 2017/2018 (0–50 scale) NPD field: KS2_GPSMARK, 2017/2018 (0–70 scale)

Participant selection

Schools were required to have sufficient numbers of EAL pupils (absolute minimum of ten in Y5 for 2016/2017, lowered to eight following slow uptake during the early recruitment period), identified using a binary EAL/not EAL measure to enable subgroup analysis. Schools sent pupil data to Enfield, which forwarded it to the evaluation team. This data was then used to personalise labels for test scripts used in the baseline measure. These were sent to schools by the evaluation team and returned by courier before being dispatched to the organisation commissioned to undertake the marking, GL Assessment.

The trial did not explicitly consider teacher eligibility, but schools were responsible for putting forward the relevant staff to take part in the training programme. These teachers were those who would be teaching Y5 classes in 2016/2017 and Y6 in 2017/2018, as per the trial timeline. We are not aware of overlap between Y5 and Y6 teaching; all participating schools were at least two-form entry, and as a result different staff taught the pupils in Y5 and Y6. Pupils joining the schools after the baseline assessment were not included in the trial as outcome testing was only undertaken with those who had done the pre-test. Soon after the baseline assessment, one school indicated that it wished to play no further part in the trial, so its 30 pupils were dropped from the analysis sample. Other schools withdrew

from the training but agreed to continue participating in the trial, and their pupils were included in the analysis sample. Further detail on this is provided below.

Outcome measures

The baseline measure was the GL Assessment Progress Test in English (PTE), a standardised assessment of technical English skills.² The Level 9 test was used, as appropriate for Y6 pupils. This was preferred to KS1 test scores as EAL pupils are known to have relatively high levels of missing KS1 data (Strand et al., 2015:8). Among our analysis sample, a higher proportion of EAL pupils (15%) are missing KS1 data compared to 4% of non-EAL pupils, suggesting that reliance on KS1 scores for baseline measures would have reduced the sample size and potentially biased the results.

Baseline tests were taken between 31 October and 25 November 2016 and administered by the evaluation team. Pupils sat the test paper in schools and completed papers were returned to the evaluators by courier service before being sent to GL for marking. Results were sent to the evaluation team from GL in electronic format, with test papers destroyed after six months. The timeframe of the set-up period of the evaluation meant that it was not feasible to use the baseline scores in randomisation. Baseline scores were used only in analysis.

The primary outcome measure was writing, assessed through a test administered by the evaluators (using KS2 past papers from 2012) for all pupils in the target year group (starting in Y5, ending in Y6). A writing test to measure subject-based literacy was sought but no suitable age-appropriate standardised test could be identified. Of the available national tests, KS2 Writing was deemed most closely aligned to the intervention aims as described in the logic model; however, changes to the marking scheme in 2016³ rendered it unsuitable for this evaluation as a more finely-graded measure (continuous rather than categorical) was needed for the power calculations to hold true. Therefore, the decision was made to administer an additional writing test in the summer of 2018. After considering the options, it was agreed to use KS2 Writing past papers from 2012, marked by an organisation contracted by the evaluation team. This was judged by evaluator, developer, and funder to provide the best solution in terms of using an outcome measure to capture pupil progress over the intervention period. As discussed in the Limitations section (page 52), the Integrating English programme meant that teachers focused in their schemes of work on writing in specific subject contexts which varied by school and sometimes by year group meaning a stronger outcome measure would focus on the specific subject context(s) covered in each particular school. Since creating a set of bespoke measures would be both prohibitively expensive and would lack comparability across the sample, it was decided to use a general test.

The test papers were marked according to four different criteria, each of which had a maximum number of marks available. These were Sentence Structure and Punctuation (0–8 marks), Text Structure and Organisation (0–8 marks), Composition and Effect (0–12 marks), and Handwriting (0–3 marks), giving a total out of 31 marks. The combined raw scores were used as the primary outcome measure. The markers were all blind to study allocation.

The primary outcome assessment was only completed in 80 of the 90 schools participating in the trial. This test was not part of the original trial design and was only added once it became clear that the NPD would not provide a measure of writing attainment in adequate detail. It is likely that some schools did not appreciate the additional burden during the summer term and did not take part for this reason, although there is no evidence for this. The evaluators made repeated attempts to contact all schools in the lead-up to the testing period. Prior to that, several notifications were sent to schools to explain the testing requirements, with the developers also informing teachers from the intervention group at training

² <https://www.gl-assessment.co.uk/products/progress-test-in-english-pte/>

³ https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/512097/2016_KS2_Assessmentandreportingarrangements__ARA__PDFa.pdf

sessions. Tests were carried out at a time agreed between the school and the evaluator so that invigilators could attend in each case. Some invigilators were supplied by an external contractor, others were recruited as temporary employees of the evaluator. The evaluation project manager personally visited five schools for this purpose. No invigilators reported any doubts over the integrity of the testing process.

The secondary outcomes were attainment in KS2 Grammar, Punctuation and Spelling (GPS) and KS2 Reading. Raw scores, as available from the National Pupil Database (NPD), were used for the impact analyses. These secondary outcome tests were marked externally as part of compulsory national KS2 testing and may therefore be seen as more methodologically robust, but are less theoretically aligned to capturing the pupil-level impact of the Integrating English intervention. All analyses are conducted using both secondary outcomes (see Analysis section for details).

Sample size

The original intention was to recruit 100 schools to the trial, with a minimum of ten EAL pupils per school. This recruitment target was calculated to support a MDES of 0.20 standard deviations for the EAL subgroup analysis, once a baseline covariate was included. This was also consistent with a MDES of 0.18 for all pupils (estimated 40 per school). However, the delivery team was only able to recruit 91 schools to the project, despite lowering the minimum number of EAL pupils from ten to eight. One school, which was randomised into the intervention group and informed of its allocation, withdrew from the trial after baseline testing and stated that it wanted no further involvement in any aspect of the project, leaving a total of 90. As a result, no outcome data was obtained for this school. Assuming the same details shown below in Table 4, this resulted in MDES estimates of 0.19 (all pupils) and 0.21 (EAL pupils).

Separate sample size calculations were not conducted for FSM pupils as this data was not available to the evaluators until after the intervention delivery was finished. Given the topic of the intervention, the priority for subgroup analysis was EAL pupils, so sample size calculations were only conducted to account for this group. Whilst an FSM subgroup analysis was undertaken, this was not determined by the Integrating English theory of change but is a standard requirement across all EEF evaluation trials as part of the EEF's specified aim of severing the link between family background and educational attainment.

Randomisation

The MinimPy programme⁴ was used to allocate schools into the intervention and control groups using minimisation with a random start. Five school-level variables were used to ensure a good baseline balance between the two groups. These were mean KS2 attainment, number of Y5 classes, percentage of EAL pupils, the number of EAL pupils in each of the categories on the fluency scale (introduced in 2016 comprising five categories of English proficiency, ranging from A, 'New to English', to E, 'Fluent'),⁵ and geographical hub area. At the school level, a good baseline balance was achieved across all five minimisation variables and two additional variables—KS1 attainment and %FSM (see SAP and Table 7 below).

Analysis was not conducted blinded to randomisation. Due to the role of the main data analyst as a project manager for the trial, it would not have been feasible for allocation to remain hidden.

Baseline data collection entailed all pupils taking the Progress Test in English Level 9. This was completed between 31 October and 25 November 2016. Allocation using minimisation took place in the

⁴ <http://minimpy.sourceforge.net/index.html>

week commencing 28 November and schools were informed of allocation on 30 November. Only pupils that completed the pre-test were considered for the randomisation.

Statistical analysis

A multilevel approach was taken to the analysis, with pupils clustered into schools as per the clustered trial design. Multilevel linear regression models were constructed for the primary and secondary outcome measures. The first model (Stage 1.1 as reported in the SAP) includes the GL PTE (Level 9) baseline (school centred) as a covariate at the pupil level. The GL PTE Level 9 was used as the baseline rather than using pupil-level KS1 data given the higher level of missing KS1 data for EAL pupils (Strand et al., 2015:8). The second model (Stage 1.2) includes both the pupil-level, school-centred GL PTE measure and school-level mean GL PTE scores as covariates. Including baseline covariates at both school and pupil levels results in lower MDES estimates than would be obtained with a single covariate at one of these levels. The model that includes prior attainment at just the pupil level is more reflective of approaches taken in other EEF evaluations and therefore the trial findings will be more directly comparable. The model that includes prior attainment at both pupil and school levels is included to reflect the original research design and power analyses for this trial.

The five school-level minimisation variables—school-level KS2 attainment, N classes, percentage EAL pupils, percentage pupils in each fluency category, and hub area—were included in Stage 2 of the analysis. The purpose of this analysis was to acknowledge the use of these variables in minimisation and to examine if and how their inclusion changed the estimated effect size for the impact of Integrating English. Including minimisation variables was intended to increase the precision of the estimates. While Stage 1.2 most closely follows the design decisions included in the power calculation, Stage 2 adds the minimisation variables to Stage 1.2. Results from Stage 2 are considered the headline results and are reported in the tables below for each analysis. In Appendix E, we report a full table containing results from all three models for the primary outcome measure to demonstrate fully the approach taken for all analyses. Additionally, this example model output is used to illustrate the steps taken to calculate the Hedges *g* effect size statistics from model coefficients.

Compliance with the intervention was defined by four criteria:

- attendance at training events;
- attainment of certificate for completing the coursework—awarded only to individuals that attended at least 80% of training sessions (both of these first two items based on data provided by the delivery team);
- the quality of Schemes of Work; and
- Project Action Plans (these two items rated on a 0–3 scale by the evaluators, where a minimum score of two was required for an individual to be classed as compliant).

This data was collected at teacher level, so school compliance was defined by whether the majority of teachers in the school were judged to be compliant. Full details can be found in the IPE section. We have no evidence that teachers or pupils in control schools received the Integrating English programme, and so assume that all control schools can be considered as compliant with the control condition.

Subgroup analysis was carried out for EAL pupils, essential given the focus of the intervention, and FSM pupils, given the remit of the EEF to investigate the effect of all interventions upon disadvantaged children. In recognition of the fact that language proficiency within the broad range of EAL pupils varies substantially, additional subgroup analyses were undertaken for EAL pupils for whom fluency scale data was available, and for pupils categorised as both FSM and EAL with fluency scale data. Complier Average Causal Effect (CACE) analysis was also conducted to examine the effects of compliance. Specifically, this entails dividing the ITT effect size and confidence intervals by the proportion of intervention pupils in the compliant schools (see Jo et al., 2008 and Schochet and Chang, 2011). Finally,

to investigate patterns in missing data on the primary outcome measure, a logistic regression model with not/missing as the outcome variable was estimated to examine variation according to the key pupil-level covariates used in the headline and exploratory analyses specified in the SAP. The missing data analysis also includes the secondary outcome scores as predictors to ascertain any relationship between attainment on KS2 Reading and GPS (combined to avoid multicollinearity) and missing data on the primary writing assessment.

Each set of analysis comprised the three models described above for the primary outcome and both secondary outcome measures. The tables presented below contain only results from the full headline models. There were six sets of analyses conducted; these are summarised in Table 3. Calculation of effect size is detailed in Appendix E.

Table 3: Summary of analysis stages as reported

Analysis	Participants	Additional controls
1	All (ITT)	None
2	EAL pupils	None
3	All (ITT)	Pupil-level fluency
4	FSM pupils	None
5	FSM pupils	Pupil-level fluency
6	CACE	None

Implementation and process evaluation

The IPE data collection was framed by an evidence-informed logic model (Coldwell and Maxwell, 2018) detailing the programme theory and indicating the intermediate outcomes, processes, and influencing factors underlying the intervention. The logic model, agreed with the delivery team in the early stages of the evaluation, is included in Appendix C. A set of methods were devised to address the IPE Research Questions (RQ) listed on page 12 of this report, as summarised in Table 4:

Table 4: Methods used for addressing each research question

RQ	Method			
	(a) Survey	(b) Training observation	(c) Case Studies	(d) Monitoring data analysis
1	✓	✓	✓	
2	✓		✓	
3	✓		✓	
4	✓	✓	✓	✓
5	✓	✓	✓	✓
6				✓

In more detail, the methods included:

(a) A pre- and post-intervention online survey, designed, piloted, and delivered to all Y5 and Y6 teachers in intervention and control group schools (see Appendix F). The pre-survey was conducted at two different times to allow for participation of both Y5 and Y6 teachers only. This meant that Y5 teachers were surveyed before randomisation and Y6 teachers surveyed afterward (as it would not have been possible to identify Y6 teachers prior to randomisation). The post-intervention survey was issued to all teachers (in Y5 and Y6) during the final half term of Summer 2017/2018. The survey intended to identify the level of support provided to teachers and school leaders when dealing with EAL students in the trial schools, their experiences of the programme, and their attitudes towards teaching. These issues were identified in the logic model (see Appendix C) as having an influence on the likely success of the intervention. The survey was designed to measure differences in relation to key topics between intervention and control schools before and after the intervention. The same tool was used for both

surveys. Topics included the teaching of grammar, the role of language across the curriculum, and the behaviour of EAL students in class. The post-intervention survey included questions on current teacher practice and changes over the previous months. A school-level response rate of 74% was achieved in the pre-intervention survey. However, the post-intervention survey obtained only ten responses from the 117 teachers who had also completed the pre-intervention survey. It was therefore decided that the data could not be used in the evaluation as there was no prospect of conducting meaningful comparisons.

(b) Observation of six training/CPD (Continuing Professional Development) events in total by evaluation team to examine the programme delivery in practice, and analysis of schemes of work. Implementation of the training programme was observed across the five hub areas, and all CPD sessions were observed in the training for either Y5 or Y6. The observers were all part of the process evaluation team and were experienced teachers and trainers. All observers had access to the LiLAC training manual and to the resources on the VLE. Observers used a prepared observation schedule (see Appendix G) aimed at finding evidence for the following key indicators of effective CPD (Maxwell et al., 2018):

- content focus of training;
- active learning in the training sessions;
- relevance to teachers; and
- claims made, and evidence provided, for the programme.

As part of the programme team's delivery model, attendance was recorded for all participants, as well as completion of inter-session tasks. Only participants with 80% attendance and task completion were awarded CPD certification. This data was used by the evaluation team as part of the measures of fidelity to the training programme.

(c) To evaluate the process of the training and its implementation in schools, the team of experienced primary, EAL, and ESL teachers and researchers that had observed the LiLAC training sessions visited schools in Y5 and Y6 to observe lessons, interview teachers and school leaders, and collect documentary evidence of implementation. Each researcher was allocated to one region.

Schools were invited to volunteer as case studies by email, and so self-selected for case study visits. A small proportion of schools in each region failed to respond to requests to take part in case studies, although there was not a consistent correspondence between the schools that responded and those that did well on measures of fidelity for the training programme. The research team visited a total of 14 schools: one school from each of the five regions following the Y5 training and two from all but one region in Y6. One region, Cambridgeshire and Norfolk, was under-represented by one school due to travel disruption for the planned school visit. At each school, researchers observed at least one lesson and interviewed the teacher of the lesson, often interviewed other teachers involved in the Integrating English project, and interviewed at least one school leader involved in the project.

For the case study visit, an observation schedule (Appendix H) was prepared with a focus on collecting evidence for language-based descriptions, for 'LiLAC' strategies and activities in class, and for evidence of student behaviour to indicate their ability to work with materials and ideas related to language in learning across the curriculum. For the purposes of this process evaluation, the evidence collected by the observation schedule served only to confirm that the Integrating English intervention was being implemented in the classroom. In each case, the researcher attended a class selected by the teacher as an example of the programme in action, and so the observer's role was to collect data that could support that view. The data was not used here to compare across classes or as a measure of differences (Hardman and Hardman, 2017). The completed observation schedules were also an aid for the post-observation teacher interview, serving as a possible focus for events in the classroom.

One interview schedule was prepared for teachers and one for school leaders (see Appendix H) for the case study visits. Researchers were also encouraged to consider follow-up questions, particularly

where they understood that interviewees felt strongly about a particular issue. The interviews were transcribed, incorporated into NVivo Pro 11, and coded deductively (Saldana, 2011) using categories from the project logic model and from the primary objectives of the Integrating English training programme utilising a Framework Analysis approach (Smith and Davies, 2010) which involves an initial overall examination of the data, developing a framework of categories (in this case deriving from the logic model), and then coding to categories and finally interpreting the data. The categories focused on the four main principles identified by the programme team as the core objectives of the training programme: the content of the training programme, resulting changes in pupils, resulting changes in teachers, resulting changes in the school, support from training programme, support from school leadership, and interaction with other training and school objectives. All interviews were coded for the four main principles of training and this data included in analysis. In addition, interviews from Y6 teachers were coded using an open frame on broader issues of relevance to the study to ensure that the wider views of participants in the trial who were the most knowledgeable and experienced with regard to the intervention and the underlying principles were included. While some interview questions directed teachers and school leaders to discuss these issues, among others, interviews were coded so that mention of any of the categories could be coded at any time of the interview, and the topic as a whole or any part of the topic could be mentioned in order to be coded. Codings were then quantified and trends were identified across schools and regions.

(d) Analysis of monitoring data on attendance at events, completion of the training, and creation and implementation of action plans, schemes of work, and shared teaching materials were also undertaken. These were combined to develop an indication of fidelity to the training programme. Attendance at training sessions and completion of training tasks constituted completion of the LiLAC training programme. This measure was recognised by the programme team by the awarding of certificates of CPD to all qualifying teachers and leaders. Action plans and schemes of work (SoWs) represented an initial measure of compliance with the implementation of the training programme that could then be corroborated through case study visits to schools. Action plans and SoWs were rated against scales of quality, designed in consultation with the programme team. The scores provided a threshold of compliance with the expectations of the training providers. The VLE for the project was monitored during and after the programme to contextualise interview comments, but as the sharing of materials was not a requirement of the training, the data was not included in the evaluation.

Costs

Cost data was provided to the evaluation team by the developers. No cost data was collected directly from schools. Costs are calculated for two delivery models. Per pupil costs are based on 60 pupils per school. This assumption is made as schools were required to be two-form entry in order to join the trial.

In calculating the costs of delivering the intervention, we present two scenarios: the school-based model where individual schools pay for external trainers, and the hub-based model where schools pool resources for teachers to be trained in group sessions. While schools choosing to use Integrating English may be unable to pool resources with others, the hub-based approach is more similar to that employed in this trial. Full details can be found in the costs section of this report.

Timeline

Table 5 (below) shows the schedule for the main intervention and evaluation activities.

Table 5: Timeline

Date	Activity
By 21 October 2016	Recruitment deadline—schools send signed MOU and pupil data to SHU
31 Oct–25 Nov 2016	Baseline testing in all schools
31 Oct 31–25 Nov 2016	Conduct pre-intervention survey for Y5 teachers
30 November 2016	Randomisation of schools to intervention/control group
Spring term 2016/2017	LiLAC training for Y5 teachers
Spring term 2016/2017	Observations of teacher CPD training
Summer term 2016/2017	Y5 pupils taught using LiLAC, Y6 teachers receive two days of training
Summer term 2016/2017	Case studies in five schools
Summer holidays 2017 / Autumn term 2017/2018	Y6 teachers complete outstanding LiLAC training where necessary
Autumn 2017/2018	Pre-intervention survey for Y6 teachers
Spring term 2017/2018	Y6 pupils taught using LiLAC
Spring term 2017/2018	Case studies in nine schools comprising school visits
November 2017	Statistical Analysis Plan agreed
Summer 2018, second half term	Post-intervention survey for all teachers
Summer 2018	Key Stage 2 testing
Summer 2018	Outcome assessment administration and marking

Impact evaluation

Participant flow including losses and exclusions

Figure 2 below shows the number of schools and pupils recruited to the trial for the randomisation stage. Before baseline testing, 91 schools agreed to take part in the trial. Data on 4,987 pupils was sent to the evaluation team. Test papers with pupil names were sent to schools, and 4,763 took the test.⁶ One school later indicated that it wished to play no further part in the trial, so its 30 pupils were dropped. It became clear that a test sheet populated with incorrect pupil codes (used by the evaluation team to identify pupils anonymously) was sent by error to one school resulting in duplicate data for one pupil. The total number of pupils in the analysis sample was 4,732. However, only 3,607 pupils from 80 schools took part in the primary outcome assessment. Further detail on attrition is provided below.

Eight schools from the intervention group indicated that they wished to withdraw from the training. Reasons included burden on staff time to attend training and changes to school management. Two of these schools later agreed to take part in the primary outcome testing. Secondary outcome data was obtained and analysed for all 90 schools.

⁶ We have calculated attrition using the number of pupils that sat the baseline test and were randomised (N = 4762).

Figure 2: Participant flow through the evaluation

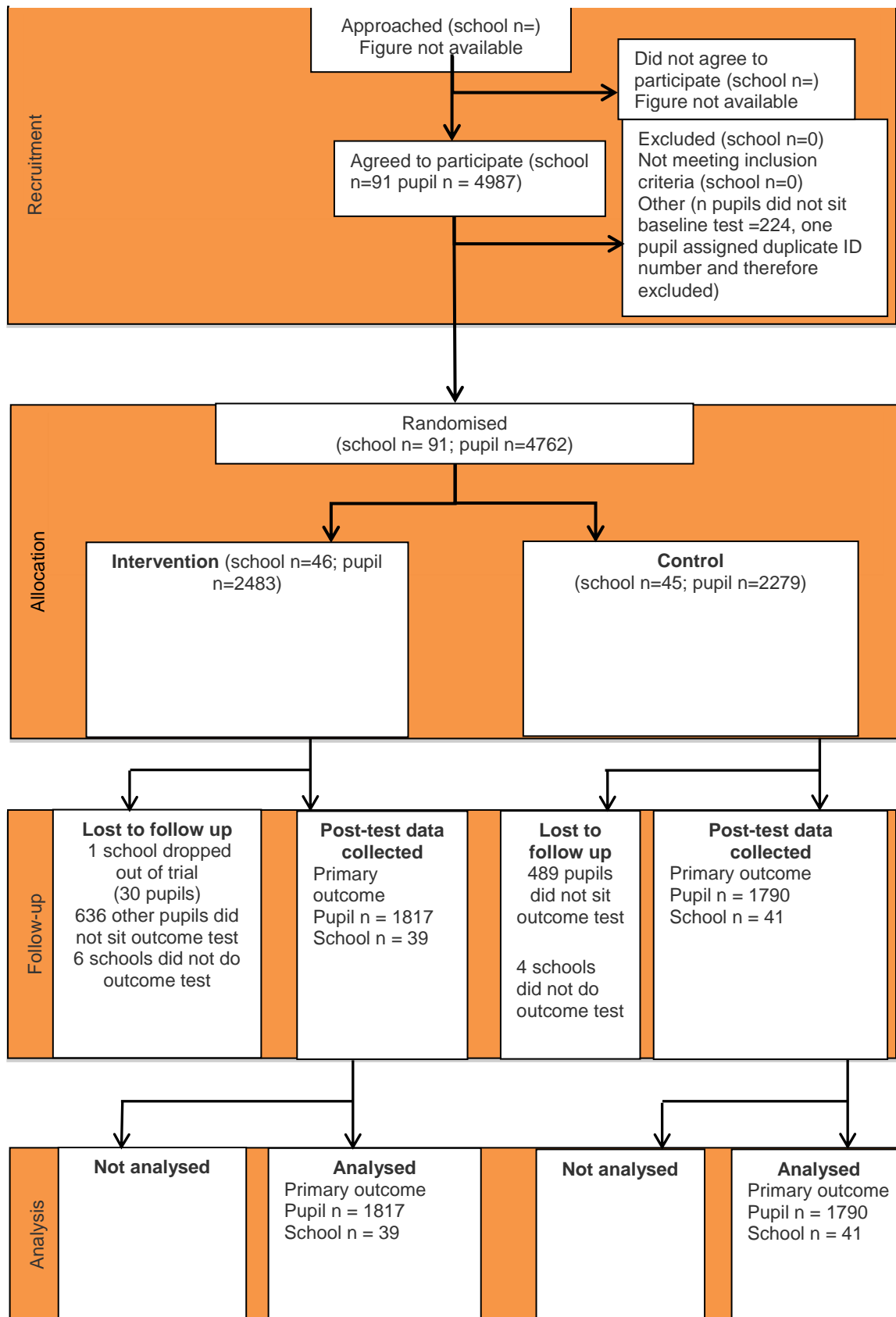


Table 6: Minimum detectable effect sizes at different stages

		Protocol		Randomisation		Analysis	
		Overall	EAL	Overall	EAL	Overall	EAL
MDES		0.18	0.20	0.19	0.21	0.23	0.27
Pre-test/post-test correlations	level 1 (pupil)	0.74	0.74	0.74	0.74	0.57	0.42
	level 2 (school)	0.60	0.60	0.60	0.60	0.60	0.60
Intracluster correlations (ICCs)	level 2 (school)	0.14	0.14	0.14	0.14	0.18	0.19
Alpha		0.05	0.05	0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		2	2	2	2	2	2
Average cluster size		40	10	40	23	45	25
Number of schools	intervention	50	50	46	46	39	39
	control	50	50	45	45	41	41
	total	100	100	91	91	80	80
Number of pupils	intervention	2000	500	2483	1298	1817	995
	control	2000	500	2279	1257	1790	983
	total	4000	1000	4762	2555	3607	1978

Attrition

From the 80 schools that did the outcome test, a total of 3,607 pupils took part out of 4,762 that completed the pre-test and were randomised. The overall rate of attrition for the primary outcome is 24%. However, among only those schools that took part in the primary outcome assessment (80 schools), the attrition rate was 15%. As might be expected, for the NPD secondary outcomes attrition was far lower. NPD data provided KS2 Reading marks for 4,531 pupils and KS2 GPS marks for 4,533 pupils. On these measures, the attrition rate is 4.85%.

Pupil and school characteristics

Table 7 below compares the intervention and control group samples at baseline (91 schools, 4,762 pupils) and in the ITT analyses for the primary KS2 writing outcome. This shows good balance at both baseline and outcome in terms of hub area, OFSTED, number of classes per school, school-level KS1 and KS2 attainment, %EAL, %FSM, Fluency in English, and pupil-level GL baseline test score. This is despite a slightly higher number of intervention schools not taking part in the primary outcome assessment.

Table 7: Comparing intervention and control samples at baseline and analysis

School level (categorical)	Baseline (N=91)		Analysis (N=80)	
	Intervention group (N=46)	Control group (N=45)	Intervention group (N=39)	Control group (N=41)
	% (n)	% (n)	% (n)	% (n)
<i>Hub Area</i>				
South East	28% (13)	22% (10)	28% (12)	22% (10)
West Midlands	24% (11)	24% (11)	24% (10)	24% (9)
The North	17% (8)	16% (7)	17% (5)	16% (6)
Leics/Peterborough	15% (7)	24% (11)	15% (6)	24% (10)
Cambs/Norfolk	15% (7)	13% (6)	15% (6)	13% (6)
<i>OFSTED Grades</i>				
Good / Outstanding	67% (31)	76% (34)	72% (28)	76% (31)
Requires Improvement / Inadequate	13% (6)	13% (6)	10% (4)	12% (5)
Missing	20% (9)	11% (5)	18% (7)	12% (5)
School level (continuous)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Number of classes	2.04 (0.594)	2.00 (0.739)	2.08 (0.579)	2.02 (0.758)
KS2 Attainment 2014/5	28.2 (1.47)	28.4 (1.58)	28.4 (1.36)	28.5 (1.56)
KS1 attainment for 2014/5 KS2 cohort	14.7 (1.32)	14.6 (1.31)	14.6 (1.35)	14.7 (1.32)
%EAL	51.3 (25.5)	53.8 (23.6)	52.2 (25.1)	54.3 (23.3)
%FSM in last 6 years	43.78 (17.13)	43.6 (19.06)	36.6 (15.5)	38.2 (17.7)
<i>Fluency in English Scale: (% pupils per school)</i>				
<i>A (New to English)</i>	2.4 (3.77)	2.3 (5.22)	2.2 (3.67)	2.6 (5.42)
<i>B (Early Acquisition)</i>	5.3 (9.64)	4.9 (8.41)	5.2 (10.22)	4.9 (8.53)
<i>C (Developing Competence)</i>	13.8 (15.51)	15.5 (15.11)	13 (15.68)	14.8 (14.46)
<i>D (Competent)</i>	15.2 (14.08)	17.3 (14.51)	16 (14.1)	17.3 (14.94)
<i>E (Fluent)</i>	13.9 (19.55)	13.4 (15.78)	15.5 (20.47)	14.2 (16.2)
<i>Fluency Details Missing</i>	0.4 (0.95)	0.6 (1.77)	0.4 (1)	0.5 (1.82)
<i>Not EAL</i>	48.9 (24.99)	46 (24.06)	47.7 (25.12)	45.7 (23.34)
Pupil level (continuous)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Pre-test score*	25.4 (10.94)	24.9 (11.48)	25.9 (10.54)	25.6 (11.22)
effect size (Int-Cont)	+0.04		+0.04	

Outcomes and analysis

As discussed above, the original plan to use KS2 Writing as the primary outcome was abandoned when the marking scheme changed in 2016. It was agreed that the best solution was to use KS2 past papers as the primary outcome. The secondary outcomes were KS2 Reading and KS2 Grammar, Punctuation and Spelling. Both secondary outcomes are still marked using continuous scales and this data was obtained from the National Pupil Database. Histograms for the outcome and baseline measures can be found in Appendix D.

The protocol and SAP specified a series of analyses to be conducted. The results are presented below. For each set of analyses, multilevel models have been estimated, with pupils clustered into schools. All analyses have been run on all three outcome variables. The results displayed relate to the full models

that contain the full set of covariates as specified in the SAP: hub area, number of classes per school, school-level KS2 attainment, and school-level percentage of pupils classified as EAL and in each fluency category. These figures represent the headline findings for the impact evaluation. As an example of the analysis process, all models for the ITT analysis are presented in Appendix E.

The first set of analysis conducted used the full intention-to-treat (ITT) sample. Table 9 shows that for all three outcome measures, no statistically significant difference between the intervention and control groups has been observed. All mean scores are very similar. The point estimates for all three outcomes are negative, but include wide confidence intervals that are consistent with negative effects and small positive effects (in the case of writing) and meaningless positive effects (for reading and GPS). Therefore, we conclude that this trial found no evidence that the Integrating English intervention resulted in pupil-level gains in KS2 Writing, Reading, or GPS attainment compared to 'business as usual'.

Table 8: ITT analysis

Outcome	Raw means				Effect size		
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	p- value
	N (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
KS2 Writing	1817 (636)	15.72 (15.45,15.99)	1790 (489)	15.79 (15.5, 16.08)	3607 (1817; 1790)	-0.05 (-0.21, +0.12)	0.577
KS2 Reading	2348 (105)	31.18 (30.77,31.58)	2183 (96)	31.61 (31.19, 32.03)	4531 (2348; 2183)	-0.07 (-0.17, +0.03)	0.176
KS2 GPS	2348 (105)	46.78 (46.2,47.36)	2185 (94)	47.59 (46.97, 48.22)	4533 (2348; 2185)	-0.08 (-0.19, +0.03)	0.153

The intervention was intended to improve language ability among all pupils but was aimed principally at EAL pupils. It was therefore agreed at protocol stage to carry out subgroup analyses for these pupils. Results from these models are displayed in Table 9. On the primary outcome measure, the mean difference between intervention and control pupils is near non-existent. On both secondary measures, slightly higher scores are observed among the control group, yet these effect sizes are very small (below 0.10 SDs) and not statistically significant at the $p < 0.05$ level. Again, the point estimates for all three outcomes are negative; the wide confidence intervals are consistent with negative effects and small positive effects (for writing) and very small positive effects (for reading and GPS).

Table 9: EAL subgroup analysis

Outcome	Raw means				Effect size		
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	p- value
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
KS2 Writing	995(298)	15.87(15.51,16.23)	983(274)	15.99(15.6,16.37)	1978(995;983)	-0.06 (- 0.25,+0.12)	0.517
KS2 Reading	1231(62)	29.89(29.32,30.45)	1197(60)	31.4(30.83,31.97)	2428(1231;1197)	-0.08 (- 0.18,+0.02)	0.134
KS2 GPS	1231(62)	46.41(45.58,47.24)	1200(57)	48.92(48.07,49.76)	2431(1231;1200)	-0.09 (- 0.2,+0.02)	0.120

Among EAL pupils a wide range of ability in English is found. It was agreed in the SAP to also undertake analyses including pupil-level fluency. The fluency scale is measured on an ordinal scale with five

categories from A (new to English) to E (fluent in English). There was missing data for 27 pupils on this variable, although ten of these did not sit the primary outcome test. Results from these models are displayed in Table 10. Variation between the intervention and control groups is negligible across all three measures. Effect sizes are all negative, smaller than ± 0.1 standard deviations and not statistically significant at the $p < 0.05$ level, as is the case for the ITT and EAL subgroup analyses discussed above.

Table 10: Exploratory analysis (including pupil-level fluency as additional control)

Outcome	Raw means				Effect size		
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	p- value
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)				
KS2 Writing	1811(631)	15.81(15.54,16.08)	1779(484)	15.79(15.5,16.08)	3590(1811;1779)	-0.05 (- 0.22,+0.12)	0.561
KS2 Reading	2339(103)	31.65(31.25,32.06)	2169(94)	31.61(31.19,32.03)	4508(2339;2169)	-0.07 (- 0.17,+0.03)	0.154
KS2 GPS	2339(103)	47.63(47.05,48.22)	2171(92)	47.59(46.97,48.22)	4510(2339;2171)	-0.08 (- 0.19,+0.03)	0.136

Exploratory analysis was conducted on the subsample of pupils recorded in the National Pupil Database as having been eligible for free school meals at any point in the past six years. The greatest variation discovered was for the secondary measure, KS2 Reading, but again, all effect sizes are all smaller than ± 0.1 standard deviations and not statistically significant at the $p < 0.05$ level. There is no evidence for any impact of the intervention on this group of pupils. Results from these models are displayed in Table 11.

Further exploratory analysis was conducted on the FSM subsample, with models incorporating pupil-level English proficiency as described above. Again, the greatest variation discovered was for the secondary measure, KS2 Reading, but effect sizes are all smaller than ± 0.1 standard deviations and not statistically significant at the $p < 0.05$ level. These figures are available from the authors upon request.

Table 11: Exploratory analysis (FSM subgroup)

Outcome	Raw means				Effect size		
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	p- value
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)				
KS2 Writing	630(271)	14.77(14.32,15.23)	661(198)	14.91(14.43,15.38)	1291(630;661)	+0.01 (- 0.18,+0.20)	0.917
KS2 Reading	883(18)	29.02(28.34,29.71)	830(29)	29.96(29.25,30.66)	1713(883;830)	-0.09 (- 0.21,+0.03)	0.129
KS2 GPS	882(19)	43.57(42.6,44.54)	831(28)	44.53(43.49,45.57)	1713(882;831)	-0.05 (-0.17; +0.03)	0.488

Fidelity to the intervention was determined by attendance at training events (minimum 80% of sessions), attainment of the certificate of completion (which was dependent upon meeting the 80% attendance criterion), assessed standard of schemes of work, and assessed standard of project action plans. Further detail on these indicators can be found in the implementation and process evaluation section of this report. The evaluation team conducted exploratory cross tabulations examining the relationship between pupil scores on each of the three outcome measures and the separate fidelity variables—the number of teachers in each school attaining the certificate of completion, SoW score, and action plan score. No discernible pattern emerged. The figures are presented in Appendix J.

Only 12 schools from the intervention group met the criteria agreed between the evaluation and delivery teams to be classed as compliant. This represents just over a quarter of the 45 intervention schools present at baseline and just under a third of the 39 intervention schools that completed the primary outcome (KS2 writing) testing. In terms of pupils, the proportion of intervention pupils located in the 12 compliant intervention schools was 29.6% for the primary outcome.⁷

To estimate the Compliers Average Causal Effect (CACE) for the primary outcome, the ITT effect size and confidence intervals were divided by the proportion of intervention pupils located in the 12 compliant schools as specified by Jo et al. (2008) and Schochet and Chang (2011). This is shown below as an equation and the effect size for the primary outcome is used to illustrate how the CACE estimate is calculated.

$$CACE\ estimate = \frac{ITT\ estimate}{proportion\ of\ pupils\ in\ compliant\ schools} = \frac{-0.05}{0.296} = -0.17\ sds$$

Table 12 summarises the CACE estimates for the primary and secondary outcomes.

Table 12: CACE analyses

Outcome	ITT Effect Size (Confidence Intervals)	Proportion of pupils in compliant schools	CACE Effect Size (Confidence Intervals)
KS2 Writing	-0.05 (-0.21; +0.12)	538/1817 = 0.296	-0.17 (-0.71; +0.41)
KS2 Reading	-0.07 (-0.17; +0.03)	606/2349 = 0.258	-0.27 (-0.66; +0.12)
KS2 GPS	-0.08 (-0.19; +0.03)	606/2349 = 0.258	-0.31 (-0.74; +0.12)

The CACE estimate focuses solely on compliance within the intervention group. We have no evidence to suggest that pupils in control schools received the intervention and therefore assume the control (business as usual) condition was maintained across the control schools. However, the pupil-level rate of compliance within the intervention group was less than 30%. This results in considerably weighting up the negative ITT effect size for the primary outcome from -0.05 to -0.17 SDs. The low rate of compliance also results in sizably increasing the width of the effect size confidence interval from 0.33 SDs in the ITT estimate (-0.21; +0.12) to 1.11 SDs for the CACE estimate (-0.71; +0.41).

Missing data

As discussed above, the main source of attrition in this trial was schools not taking part in the additional assessment which provided primary outcome data. Of the ten schools that did not participate in the primary outcome test, six were from the intervention group and four from the control group. This school-level attrition for the primary outcome did not cause substantial imbalance in the analysis sample as can be seen in Table 7.

Table 13 shows that once all pupils at schools that did not sit the primary outcome test are removed from the sample, pupils in intervention schools are no more likely to have missing data for the primary outcome measure (odds ratio = 1.03; CI 0.86, 1.24). Pupil scores on the baseline test scores did not determine the likelihood of missing the primary outcome assessment (odds ratio = 0.99; CI 0.97, 1.00; non-significant). The secondary outcome measures, combined into a total score here to circumvent multicollinearity, are negatively associated with a missing value for the writing test, but the relationship is very weak (odds ratio = 0.99; CI 0.99, 1.00; $p < 0.001$). Looking at pupil-level fluency, all fluency categories are more likely to have not taken the primary outcome assessment compared to the reference group (A: New to English, the weakest level of proficiency), with the exception of the second lowest level (B: Early Acquisition; odds ratio 0.94; CI 0.41, 2.16; not significant), although the confidence

⁷ For the primary outcome (KS2 Writing, past papers) 538 pupils were located in the 12 compliant schools out of a total of 1,817 pupils across the 39 schools that participated in the testing for this outcome: $538 / 1817 = 0.296$.

intervals are very wide for each category. Of these estimates, non-EAL pupils are also more likely to have missed the outcome assessment compared to those in the reference category (odds ratio 1.73; CI 0.84, 3.55; not significant), but again the confidence intervals are wide. Pupils classified as FSM are more likely to have missing data for the primary outcome assessment (odds ratio 1.35; CI 1.12, 1.64; $p < 0.05$).

Table 13: Missing data analysis

Missing	Odds Ratio	Std. Err.	Sig.	95% CI lower	95% CI upper
Intervention	1.03	0.099	0.75	0.86	1.24
Baseline score (centred)	0.99	0.007	0.08	0.97	1.00
KS2 (Reading and GPS) combined	0.99	0.003	0.04	0.99	1.00
Fluency (reference A)					
B	0.94	0.400	0.88	0.41	2.16
C	1.49	0.566	0.30	0.71	3.14
D	1.82	0.689	0.11	0.87	3.82
E	1.35	0.524	0.43	0.64	2.89
Non EAL	1.73	0.635	0.13	0.84	3.55
FSM	1.35	0.133	0.00	1.12	1.64
Constant	0.13	0.053	0.00	0.05	0.29

N = 4,034.

Within the Statistical Analysis Plan, we specified that we would examine the extent to which missing data could alter the estimated impact of Integrating English on the primary and secondary outcomes by drawing on Gorard (2015) and Gorard and Gorard (2016) to examine the number of cases needed to disturb a finding. Specifically, this would be to calculate the number of cases needed in order to reduce an observed effect size to zero. However, given that the ITT impact analyses concluded that the impact of Integrating English was not significantly different from zero for primary and secondary outcomes, in this trial there is no 'finding' that might be disturbed through the inclusion of missing cases. It would be possible to calculate the number of cases needed in order to move the (not significantly different from zero) ITT effect size to precisely 'zero'; but this would bring very limited additional insight other than underlining that we found no evidence that Integrating English led to pupil gains in KS2 attainment (writing, reading or GPS).

Cost

In calculating the costs of delivering the intervention, we present two scenarios. The school-based model involves schools independently hiring trainers to attend and train their staff. Schools are able to make these arrangements independently and this therefore reflects the costs that would be incurred should schools choose to use the programme without partnering with other schools to pool costs.

However, during this trial, training was delivered through geographical hubs. On average, these trained 20 teachers each. Should schools wish to run this intervention in the future, the cost of hiring trainers can therefore be shared between schools, reducing the cost of the intervention. Additional costs could come from venue hire, teacher travel to training venues, or stationery, all of which would be avoided if training were delivered through schools. As these potential extra costs are difficult to estimate accurately and ultimately avoidable if training were to be delivered through schools, they are excluded from our calculations. We present cost figures for this hub-based model as it is most similar to that adopted in this trial, assuming that ten schools would each put forward two teachers for the training. It is possible that Local Authorities, Multi-Academy Trusts or other groups of schools may decide to use Integrating English on this basis.

Integrating English entails set-up costs to schools, but no additional running costs after the initial outlay. The largest costs are trainer daily fees, which as mentioned could be shared between schools under a hub-based model. In the trial, schools were required to be two-form entry not only to ensure sufficient pupil numbers, but also, as the developers argued, that teachers would benefit from having a colleague with whom their training and preparation could be discussed. We therefore calculate costs assuming that two teachers per school will take part. For staff training and supported lesson preparation—both integral elements of the intervention—six days of supply cover are required per participating teacher. This totals twelve days of cover per school. Payments to trainers and staff cover are one-off costs required only during the first year. Once trained, teachers are able to deliver the intervention each year at no further cost.

The only other expense is the LiLAC manual. It may be possible for schools to share copies between staff. The manuals are ordered from Australia, so prices in GBP are subject to currency fluctuations. At the time of writing (April 2019), the AUD 180 converts to £98. Again, this is a one-off expenditure.

The intervention is delivered in normal classroom time. It is expected that teachers incorporate the principles learnt in the training into their everyday teaching. There are no ongoing costs associated with delivery. The cost per pupil, assuming 60 pupils per cohort, is £62 under the school-based delivery model and £8 using the hub-based approach (based on ten schools with two teachers per school). Calculating costs for pupils over three years involves simply dividing these figures by three, as the intervention is delivered over one school year and three cohorts of pupils can be taught for the same set-up costs. Over three years, the programme can be delivered for £21 per pupil under the school-based model and under £3 per pupil for the hub-based model.

Table 14: Cost of delivering project

	School-based model			Hub-based model		
	Unit cost (£)	N	Total (£)	Unit cost (£)	N	Total (£)
Direct marginal costs						
Tutor daily fee per school	600	6	3600	600 (/10)	6	360
LiLAC Manual per school	98	1	98	98	1	98
Prerequisites						
Staff time						
Cover for teacher training/supported preparation time		12 days			12 days	
Total			3698			458
Per pupil (60 per school)			61.63			7.63
Per pupil over 3 years			20.54			2.54

Table 15: Cumulative costs of project

Model	Year 1	Year 2	Year 3
School-based	£3698	£3698	£3698
Hub-based	£458	£458	£458

Implementation and process evaluation

This section contains the key findings from the IPE evaluation. The results of the process evaluation are presented in detail, following a key summary in response to the research questions.

Summary: responses to the research questions

RQ1 How effective is the Integrating English programme in developing teachers' knowledge and understanding of language, based on systemic functional linguistics, in different subjects?

There is evidence from interviews and observations (see 'Evaluation of Training' below) that the Integrating English delivery team effectively introduced participants to the main principles of the LiLAC training programme as well as providing the teachers with strategies and teaching techniques that were relevant to their classrooms. The fact that all case study teachers mentioned genre-based teaching as something they valued when discussing the merits of the intervention with the developers, for instance, indicates effective CPD (see 'Fidelity Measures: Implementation in Case Study Schools' below). Unsurprisingly this effectiveness was not even—only around 70% of the teachers mentioned structured talk (talk-as-performance or talk-as-process). Most observed classes included examples of language-based tasks or discussions, but the evidence of linguistic knowledge being used in class was limited to isolated incidents and a narrow range of genres, partly because not all teachers were confident in their ability to introduce knowledge about language into the classroom.

In interviews, teachers talked about a range of different techniques and strategies from the Integrating English training, and some discussed how these had affected their knowledge of teaching. Some teachers reported that the training refreshed their knowledge and practice, for others it renewed or deepened their existing understanding, and for others their teaching knowledge and practice was enhanced as a result of their engagement with Integrating English.

RQ2 What evidence is there that this knowledge and understanding leads to improved classroom practice?

There was widespread praise for the training programme from case study school interviewees (see 'Fidelity Measures: Implementation in Case Study Schools' below), and the majority of these teachers and school leaders described how easily the aims of the training integrated with wider aims of the school, other initiatives and training programmes, as well as current practice. From the pre-intervention survey it is clear that teachers from intervention and control schools recognised the central role of language across the curriculum, and so Integrating English was not the first encounter most of the teachers and school leaders had had with many of the underlying principles of LiLAC. Teachers and leaders from different schools were able to name at least three other programmes that they perceived as having similar means or aims.

A minority of teachers saw the ideas in Integrating English as hardly different from their current practice, while most emphasised the positive effect the programme had on their ability to respond to the needs of their EAL students. A few techniques such as dictogloss were widely adopted, while there was little evidence from interviews, observations, or schemes of work that the whole of the teaching–learning cycle was enacted in most classrooms (see 'Fidelity Measures: Implementation in Case Study Schools' below). Most teachers reported using new practices in the classroom, such as including a talk-as-process stage or offering planning tools for writing. Many teachers drew on descriptions of different genres in their classes, referred regularly to 'register', or enabled students to build complex nominal groups. These all represent techniques and strategies introduced by the Integrating English project to project classrooms that were often taken up enthusiastically by the teachers in the treatment group.

RQ3 What evidence is there that this knowledge and understanding leads to improved pupils' language performance in the classroom?

Although the process evaluation did not gather direct evidence of pupils' language performance, comments by teachers and school leaders in most schools (see 'Fidelity Measures: Implementation in Case Study Schools' below) offer indirect evidence for improvements. One school leader emphasised, for example, that writing outcomes across different subjects had recognisably improved during the intervention. One of the teachers described how colour-coding provided pupils with a visualisation of cohesion in text, so that pupils could discuss the choice of pronouns and conjunctions. The planning time and tools introduced resulted in improved written and spoken outcomes, particularly for EAL students, according to many teachers. Finally, for both EAL and other students it was commonly reported that confidence was greatly improved, with some EAL students gaining the confidence to work on tasks alongside other students.

RQ4 What issues of fidelity occurred during the trial?

See 'Fidelity' section below. With 85–88% of participants being awarded a CPD certificate, the training programme showed a high level of completion. However, the measures of fidelity in the process evaluation suggest that there were only a minority of schools (31%) that were fully compliant with the demands of the programme. Even removing the requirement for an action plan, there was still barely 50% compliance for schemes of work submitted to the Virtual Learning Environment, Fronter. While these demands for full compliance—CPD completion, action plans, and schemes of work—may be considered stringent, without evidence of implementation in classes with institutional support, it is likely that the intervention is not being fully implemented and that the strong support for Integrating English expressed by the teachers is for elements of the LiLAC programme rather than a full implementation of the programme.

RQ5 What does the trial indicate about scalability?

The question of scalability is not addressed in full here as the impact evaluation showed no significant improvement in scores for pupils in the intervention schools compared to control schools. It is presumed that the programme will not be scaled up in its current form. Rather than looking at scaling the intervention, we suggest a focus on simplification and narrowing focus which could be tested and then, if successful, be taken to scale, as discussed in the concluding section.

RQ6 What are the intervention costs?

The section on 'Cost' details the overall programme costs. Few comments were made in relation to this aspect of the programme, but the costs of replacing year group teachers for five days of training were characterised in a few schools as being onerous due to the disruption to teaching and the inevitable 'catching up' required of the class teachers on their return from training.

Implementation summary

The Integrating English programme consists of a number of elements. The main training programme is based on the 'Language in Learning Across the Curriculum' (LiLAC) course designed by Custance, Dare and Polias (2012) and published by the South Australia Government, including readings and between-session activities, which are discussed in the following sessions and/or on the Virtual Learning Environment, Fronter. Participants who successfully attend at least 80% of sessions and complete tasks are rewarded with a certificate of completion for CPD. The delivery team expect trainees to understand the key principles of LiLAC. As a result of the training, school leaders are responsible for submitting an action plan that outlines how LiLAC principles will be implemented in the school, and each school is expected to submit schemes of work (SoWs) that detail which Integrating English strategies will be

implemented when and in which part of the curriculum, as mediated through the Integrating English programme.

Evaluation of training

The first part of the evaluation of the programme relates to the quality of the training sessions and the required outcomes of the CPD programme. This is evaluated through observations made by the members of the evaluation team as each visited the regional training centres, making a total of six visits over the two rounds of training. The effect of the training is then assessed, in this section, by reviewing the quality of the action plans and the schemes of work, which were produced by each school as a result of the training programme. Indicators of compliance with the training programme will be discussed in further detail in the 'Measures of Fidelity' section.

Observations of training sessions

Using an observation tool (Appendix G) designed to identify indicators of effective CPD (Maxwell et al., 2008), observers found evidence of a clear focus on the content of the training in all five training centres. Content included introducing new concepts, particularly around register and linguistic realisations of generic features, and teaching techniques, such as dictogloss, as well as integrating topics into teachers' current concerns, including the need to cater to the diversity of students in the classroom. Although local factors, such as an extreme weather warning, occasionally conspired to shorten CPD sessions, all centres were clearly focused on covering the full content of the Integrating English programme.

Participants were regularly engaged in the activities and techniques they were being introduced to across all training centres; they learnt by doing. Analysis and application activities were common throughout the CPD sessions and all teachers were required to practice short tasks in their own classrooms between training sessions. The participants were asked to reflect on these and the set readings and share their comments in the following training session or online on the VLE, or both. Training sessions included discussion of reflections and comments. Participants were required to complete these tasks in order to receive their certificate of completion which signified full compliance in training for the purpose of this evaluation.

Observers noticed repeated attempts to make the training relevant to the concerns of the participants. In particular, as intended by the Integrating English programme, this required adapting the LiLAC training manual and materials from the Australian to the British educational context. The training days included a combination of new, particularly linguistic, concepts and practical classroom techniques and activities. The focus in all cases was on how these ideas would benefit both EAL and English first-language students, especially those from disadvantaged backgrounds, and time was spent providing participants with practical experience through a range of activities.

Participants' experience of implementing activities in their own classes between training days contributed greatly to their being able to evaluate first-hand the claims made by the training programme, in addition to the frequent references by trainers to the many successes of closing the gap between disadvantaged and EAL students in the Australian context. This perception by observers was supported by comments provided by participants during interview in school visits:

The training was brilliant, the days were brilliant, because I think there was enough balance of them giving talking, then getting us to do, then also giving us time as a school to talk about how we're going to implement all the strategies they're going to give us. So it was a nice balance (Y6 teacher, West Midlands).

The final training day included support in preparing schemes of work to be implemented in participants' classes after the training. The model SoWs shared by the trainers were quickly and willingly adapted by the teachers to better match their own contexts, as witnessed in the training session and as described

later during school visits. This provided a clear indication that the overall goal of the training programme had been effective as teachers were able to directly apply the intervention principles to their own context.

Overall, observers returned from training days with praise for the professionalism, relevance, and supportiveness of the trainers and the programme. In one centre, as discovered later from interviews and from discussions with the programme team, there had been some tension caused by comments made by a trainer about another trainer. Although the issue was resolved, it is likely that the incident affected some of the participants in that region.

Quality of action plans

Schools that participated in the training programme were expected to submit an action plan. As each school was encouraged to send a representative of the school leadership team, it can be safely assumed that this was one of the main outcomes of the training programme for these participants. Schools were expected to submit the action plans to the project team, and they were hosted on the VLE. The action plans were scored by the evaluation team according to the criteria below:

0. Action plan not submitted.
1. Plan recounts stages of LiLAC training but provides no strategic planning.
2. Plan focuses on achievable stages to implement training.
3. Plan details how to achieve outcomes affecting school, staff and students.

To qualify as compliant, schools need to have action plans adjudged to be at level 2 or above. Action plans were reviewed at the end of the project and so may have been submitted at any time. Plans could be submitted for Y5, Y6, or both.

Of the 39 schools that attended and completed the training programme, 21 submitted an action plan that was judged to be compliant. Seven schools submitted action plans that were not compliant, and 11 failed to submit any action plans. These results are discussed further in relation to fidelity to the training programme (see 'Measures of Fidelity').

Action plans that were submitted but considered non-compliant mentioned Integrating English strategies or techniques, but did not indicate how these would be integrated with the running of the school or the planning of other learning. Nine of 21 compliant schools submitted action plans judged to provide details of how Integrating English outcomes would be achieved and how the school, teachers and students would be affected. The best examples among these provided details of where Integrating English would be implemented in the curriculum, who would be responsible, what resources would be required, what milestones were expected and when they would be completed, and how students would benefit. Most of the action plans judged at level 3 followed the same document template provided by the programme team. Many of those that complied at level 2 used different formats, often carrying a school logo, suggesting that even though the format made implementation harder to evaluate, the implementation was itself partly embedded in school practices.

Quality of schemes of work

Schools completed and submitted, to the VLE, schemes of work for Y5 and Y6 classes. Each school was expected to plan three LiLAC-rich lessons per week across different subjects within the curriculum for the Spring term of Y6 to be fully compliant with the training programme. The criteria below, based on the requirements of the programme team, were used by the evaluation team to assess the SoWs by focusing on the quantity and range of activities or principles, and their integration within the curriculum:

0. No samples provided for the period.
1. Examples of LiLAC insufficient (less than twice a week) or show insufficient integration of key principles.

2. Examples of at least two LiLAC activities and/or principles per week in evidence, but could be irregular, unclear or incomplete at times (within the limits of the format of the document).
3. Regular integration of LiLAC activities and principles.

Of the 39 schools that had committed to the training and implementation, 15 failed to submit SoWs for the specified term (Spring, 2018). However, the vast majority (36 out of 39) submitted SoWs at some point in Y5 or Y6 of the programme. That is, overall, most schools in the project completed SoWs, but this was not implemented consistently, leaving over 50% of the schools non-compliant in the specified term. To qualify as compliant, schools need to have SoWs adjudged to be at level 2 or above (see criteria above). The 'Measures of Fidelity' section below details the level of compliance. Here the focus is on how well the SoWs were completed. This review focuses on the evaluation of the SoWs from the 24 schools that submitted for the Spring term of Y6. Five of these schools submitted SoWs that were judged to be non-compliant.

To aid analysis, most schools highlighted LiLAC activities or lessons, often using a lilac colour scheme to signal the implementation, and in general it was clear to see which part of the curriculum schools focused the programme on. It must be remembered, however, that SoWs are normally institution-specific and no two documents were identically formatted.

The five SoWs that were judged to be non-compliant typically failed on the criteria of quantity because they failed to average over two LiLAC activities or lessons per week. Care was taken not to only use the school's method of signalling LiLAC activities or lessons and it was understood that other school activities could take priority some weeks making it difficult to implement the programme at times. For this reason, an average count per week was used when calculating the quantity of LiLAC activities. In a small number of cases, SoWs were considered non-compliant because the LiLAC activities were very limited in scope (for example a familiar technique, such as dictogloss, was used repeatedly without reference to other aspects of the programme). This occurred more frequently in initial SoWs, particularly in the Y5 group, but there was also evidence of this in Spring term Y6 SoWs.

The compliant SoWs were evenly split between those scoring two and three on the criteria above. Some SoWs fell short of the highest standards because LiLAC strategies were implemented unevenly, so that on average the SoWs described a sufficient number of activities but these may have been concentrated in some weeks and noticeably omitted in others. Alternatively, some SoWs provided a sufficient quantity of activities but these were either focused in just one area of the curriculum or demonstrated a narrow range of LiLAC activities and principles (that is, no more than three).

The best SoWs integrated language awareness (such as nominalisation), strategies (the register continuum), activities (inference mat), and stages in the learning cycle (talk-as-performance or joint deconstruction) across curriculum subjects and through the term into a coherent and mutually supporting programme of activities that combined with curriculum objectives. This is likely to have been the result of intensive planning and discussion, facilitated by the allocation of resources by the school leadership team, as described in a number of interviews:

I think time was the important thing. So we knew that, particularly the first plan, when we were putting that together, [name of school leader] gave us some time to sit together and do that, which we wouldn't have had otherwise, so that was useful (South East, Y6 teacher).

Other participants noted that providing time within the training programme for planning SoWs was also appreciated:

Also with the planning, the supportive nature of the course in the way that it was actually structured as well, so that there were some sessions towards the end where we could sit with the course leaders and sit and plan and say, 'Well what about this idea? Do you think that would work? How do you think that might pan out?' We found that quite invaluable, that sacred time of two hours, just to be able to sit and

look through things to really get our heads around things that we might want to do with what we'd been shown (South East, Y6 teachers).

The SoWs provide evidence of the training programme being implemented in the planning of lessons and objectives across the curriculum to varying degrees. From the observations and interviews during school visits, there was no evidence that these were only exercises 'on paper'. There appeared to be no discrepancy between what was outlined in the SoW and the lessons observed in schools, although events in class may, as always, divert attention away from the main learning objective for a host of reasons.

Barriers to implementation

A number of comments were made in teacher and school-leader interviews regarding the length of the training. A small number of teachers and leaders made the case that providing cover for Y5 or Y6 teachers for five full days presented the school with logistical challenges and the teachers with pedagogical dilemmas:

So I would say in the first year I covered everyone who was going on the course in terms of trying to sort the logistics of that and that can be quite full-on, to take a Y5 teacher, or a Y6 teacher, out for five sessions over quite a short period of time when you think there are probably other things that they will need to go on as well (Y6 School Leader, South East).

In most cases, interviewees felt that the training programme was comprehensive and efficient, suggesting that no time was wasted, with only one school suggesting that the content of the course did not merit five days of training. This school largely offered negative comments about the training programme.

In interviews during school visits a number of participants, while generally complimentary about the training sessions, pointed out that the between-session work placed a burden on them. The teachers typically were required to catch up on what had been covered while they were away, plan, implement in class, and report on a strategy from a previous session, and read and respond to articles. Many considered this too much for a practising teacher and this was given as the reason that some school leaders did not achieve their CPD certification.

A number of comments mentioned that the VLE, Fronter, intended as the online discussion and file-sharing platform for the intervention, was not user-friendly. Some participants lost their confidence in the platform during the training sessions as uploading documents proved problematic and some complained that their contributions were somehow lost on the platform. One Y6 teacher (Leicester/Peterborough) claimed, 'Fronter is a nightmare. I've not yet met any teacher that likes it.' As the main source of evidence for action plans and SoWs, it is possible that the VLE was a barrier to providing evidence of implementation (although the delivery team offered the alternative of email and reminded schools of this). Schools that did not offer SoWs were contacted by email by the programme team after the end of the intervention period, but this did not produce any new evidence.

Teachers and school leaders were asked directly to comment on how the aims of the Integrating English programme impeded, or were impeded by, other school aims, objectives, and priorities. Very few of the interviewees suggested any conflict with other school priorities. On the contrary, teachers and leaders replied that the Integrating English programme contributed to, variously: reading, oracy, writing and general literacy aims, topic-related objectives, reducing the gap between different groups of students, and wider school objectives. Other training and school-level initiatives that were seen as complementary included 'Success for All' and 'Talk for Writing' which was described as offering the 'same', 'similar', 'superior', and 'inferior' support to students' development in writing. A few interviewees discussed the need to constantly balance priorities and make space for new ideas, such as Integrating English, where they were considered beneficial. Examples of school priorities that did not match the Integrating English

programme were identified as the different grammatical terminology used in the GPS test, Y6 SAT exams in general, and where 'Success for All' was seen as inflexible.

Participants readily acknowledged the support provided to maintain the programme. Some teachers highlighted the ease with which they were able to contact the Integrating English trainers, others described how the Integrating English project had affected the school objectives, with the full support of school leaders, while others made clear the centrality of school leadership in allocating important resources to ensuring the success of the programme. School leaders provided support by releasing teachers and leaders for training, for planning, for sharing ideas, and for providing in-school training, and by offering their vocal support to the project in many cases. It is likely that all of these elements played a key role in all of the schools that we visited where they reported a noticeable change.

Fidelity

The section below outlines the fidelity measures used for this project. After reviewing measures of fidelity from the training programme, evidence of implementation in case study schools, based on data from classroom observations and interviews with school teachers and leaders, are reviewed. Finally, reflecting the expected outcomes of the logic model (Appendix C), the effects of the intervention on teachers, on pupils (as perceived by teachers and school leaders) and on school practices are summarised, drawing again mainly on interview data.

Relatively few schools (12 out of the 45 originally designated as intervention schools) were able to fully comply with all of the demands of the evaluation—to attend all training events, complete all tasks, and submit an action plan and SoWs a term after the final training session. It is possible that the project demanded too many instances of evidence of implementation or—as the evidence from SoWs, observations, and interviews also suggests—that few schools fully implemented the training as intended by the programme.

Fidelity measures: compliance in training programme and implementation

Measures of fidelity

To assess the fidelity of the intervention across schools, participants' attendance at training sessions and completion of set reading and classroom practice tasks between sessions were registered, and documentation was produced to evidence the schools' engagement with the programme after the training had been completed. Documentation included submission of Action Plans that demonstrated how Integrating English would be implemented by the school and submission of Schemes of Work that detailed the use of Integrating English principles and activities in the class.

Attendance

For Y5 training, of 103 participants from 44 institutions in the training programme run by the Enfield team in five regional centres, 88 attended the required number of sessions (missing no more than one training day) and completed at least seven of the eight required modules to be awarded a certificate of completion for the training. Three participants were not awarded the certificate as they did not fulfil the requirements and 17 belonged to six schools that that did not complete the training as an institution. One participant who did not complete the training was a visitor from a local authority. The two remaining non-completing participants were from different schools, and in each of these schools there were sufficient teachers that completed the training to ensure that the project could continue. In short, sufficient teachers from 37 of 43 schools completed the requirements of the training programme.

In the training of Y6 teachers, there were 89 participants from 37 institutions in the same five regional centres. No schools officially withdrew during Y6 training, so the overall completion rate (88%) was higher than in the training for Y5 teachers (85%), although the number of participants not completing

all components (attendance and tasks between sessions) was higher (11) compared to the Y5 training programme (3). Two schools received training for Y6 teachers even though the school had withdrawn from the trial in Y5. Two participants that did not complete all components attended the training but were not representatives of schools participating in the trial. In all but two cases, there were sufficient representatives from each school to ensure the training could be implemented. Thus, two schools effectively, if not formally, withdrew from the trial by not having sufficient teachers complete the training. These two schools also failed to submit schemes of work for Y6 classes. Four schools that completed the Y5 training did not send any teachers to Y6 training.

Overall, attendance at training sessions and the successful completion of reading and reflection tasks between training sessions were completed by Y5 and Y6 teachers in 31 schools. Training was completed only in Y5 in six schools and only in Y6 in two schools which had withdrawn in Y5. The summary in Table 16 reveals that in the 31 schools that successfully completed training in Y5 and Y6, there were 154 teachers across the two years.

Table 16: Total number of schools and teachers completing training for Y5 and Y6

	Completed Y5 training	Completed Y6 training	Completed both Y5 and Y6 training
Schools	37	32	31
Teachers	88	78	154

Completion of action plan and schemes of work

All participating schools were expected to complete a strategic action plan and schemes of work, as described above.. These were mostly submitted through the VLE. Action plans were judged compliant if there was evidence of strategies or stages of implementation, or verifiable outcomes, and SoWs were compliant if there was evidence that at least two LiLAC interventions were planned for each week (see Quality of Action Plans and Quality of Schemes of Work above). Table 17 summarises the level of compliance for these two requirements.

Overall, a few more schools were more compliant in preparing action plans, typically completed by school leaders, than they were in completing SoWs. While the training programme included one session aimed primarily at preparing SoWs for the same term, the compliance measure was assessed one term after the training for the second cohort (in the Spring term of Y6). That is, these schemes of work were prepared independently of the training programme and needed to be uploaded to the project VLE by the school. For a school to be fully compliant, both action plans and SoWs for Spring term 2018 had to be completed, uploaded, and assessed as compliant. This resulted in a total of 12 (31%) of the schools being judged as fully compliant. These schools were concentrated in one region; 42% of the fully compliant schools came from a region that contained 31% of the schools in the project.

A higher rate of compliance may have been achieved with a different date of measurement, perhaps including Y5 SoWs where some of the non-compliant schools did submit SoWs. Comments during interviews suggested that many teachers struggled with Fronter as a technology, and this may have prevented some schools from uploading documentation. Anecdotal evidence from a small number of schools also suggested that when a school leader was moved, left the school, or was absent for an extended period, the projects that they were associated with, such as Integrating English, were rarely handed over to another school leader. This can lead to documents being left incomplete, lost, or simply not uploaded.

Table 17: Number of compliant schools for schemes of work and action plans by hub

Hub	Action Plan Compliant	SoW Compliant	Fully Compliant
South East (12 schools)	6	9	5
West Midlands (9)	6	4	3
North East (5)	3	1	1
Cambridgeshire / Norfolk (7)	3	3	2
Leicester / Peterborough (6)	3	2	1
Total	21 (54%)	19 (49%)	12 (31%)

Fidelity measures: implementation in case study schools

Implementation of training

The second approach to evaluating fidelity to the intervention is to establish to what extent the training influenced the teachers and school leaders through observation schedules (see Appendix I and following section) and interview questions (Appendix H) designed to indirectly elicit the key pedagogic approaches, which can be summarised by four principles of the LiLAC training programme, according to the programme team:

- genre-based pedagogy;
- the register continuum;
- structured talk ('talk-as-performance' and 'talk-as-process'); and
- the teaching and learning cycle.

Four principles of LiLAC training

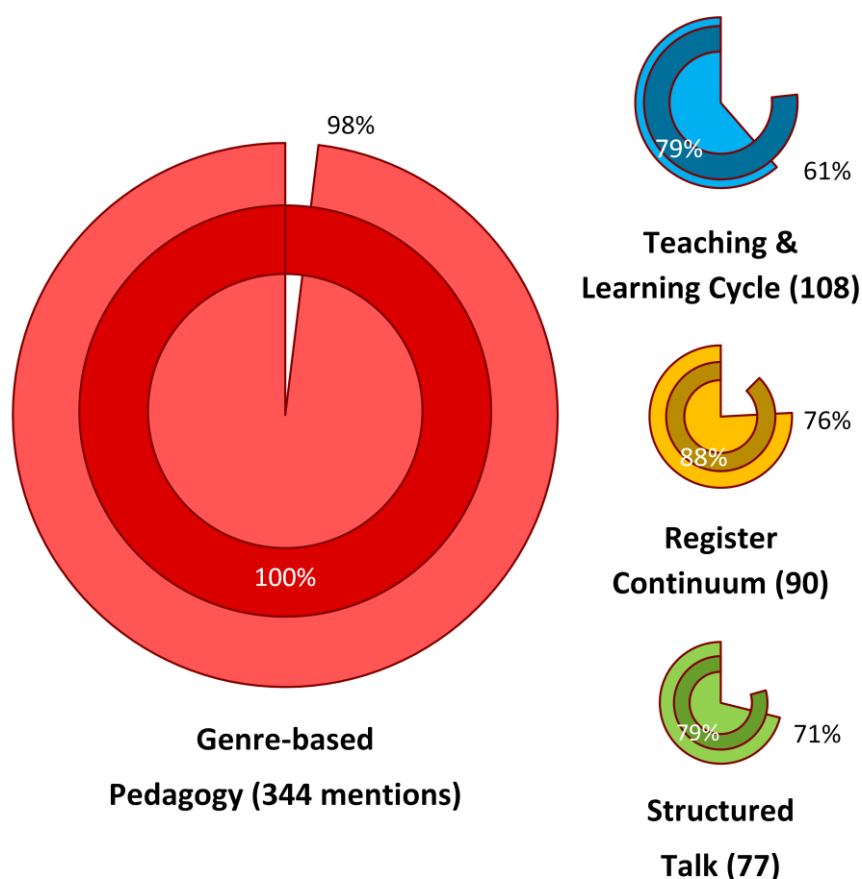
The programme team emphasised the importance of these four main principles in the training programme. That is, while the Integrating English programme aims at a wide range of skills and knowledge that target improving subject-based literacy, particularly for EAL pupils, these four central principles support the programme as whole. Consequently, evidence of the principles being implemented in classrooms was taken as indicative of the effectiveness of the training programme. Interviews from case study schools were analysed deductively by looking for evidence of the above four key principles. The naming of these principles, or aspects or components of any of these concepts, was counted as a 'mention' for the topic. Table 18 quantifies these mentions.

Table 18: Number of mentions of four principles of LiLAC training by teachers and leaders in Y5 and Y6 interviews

Year of training and Position		Genre-based pedagogy		Register continuum		Structured talk		Teaching and learning cycle	
		n	mentions	n	mentions	n	mentions	n	mentions
Y5	Leaders	5	29	4	7	2	6	2	6
	Teachers	6	91	6	12	6	23	5	30
Y6	Leaders	11	55	6	14	8	12	4	5
	Teachers	18	169	15	57	13	36	14	67
Total		40	344	31	90	29	77	25	108

Figure 3 illustrates the data in Table 18 with the size of each circle in proportion to the total number of mentions, the outer circle showing the proportion of all interviews that mentioned that topic and the inner circle representing only the proportion of teachers interviewed that mentioned the topic. Again, the size of the inner circle is proportional to the number of mentions by teachers. In this graphic it is clear that only genre-based pedagogy, or some part of it, was mentioned by all 24 of the teachers interviewed when asked which ideas from Integrating English training were used in the observed class, were of value to students, or had been used in lessons. Most mentions of the teaching and learning cycle were by the teachers, rather than school leaders, and structured talk was the least discussed principle, although almost as many teachers mentioned this topic as the register continuum, which received more repeated mentions.

Figure 3: Mentions of four LiLAC training principles in case study interviews with teachers and school leaders



Key:

Outer circle: proportion of all interviews that mentioned key principle

Inner circle: proportion of teacher interviews that mentioned key principle

Size of circle: scaled to total number of mentions

Genre-based teaching

In 40 of 41 interviews, genre-based teaching and learning was mentioned, making it the most salient key principle to be remembered after the training by teachers and school leaders. In 11 interviews it constituted over 5% of the comments, and in 17 interviews it was mentioned more than ten times (averaging 8.6 mentions). The one interviewee that did not mention this aspect of the training represented the one school that we visited that was largely negative towards the training programme.

'Genre-based teaching' focuses on the varying language features of different types of text and ways to make that knowledge accessible to students. Different ways of understanding and implementing genre-based Teaching were provided by teachers and leaders. Some described a genre:

We talked about what a non-chronological report is and what they think it is and we dissected that (Y6 teacher, North East).

Others were able to describe the effect of understanding genres across the curriculum—a key objective of Integrating English:

What LiLAC has done is it's focused us on outcomes—working from the outcomes working backwards and putting in the scaffolds and the resources that the children need to be able to write consistently across the subjects (Y5 teacher, East Midlands).

All of it is all about language anyway, so we've got to try and make sure that we're not just narrowing it down to an English lesson or we're just going to teach English in English. We've got to make sure that we teach it across the curriculum (Y6 school leader, East Midlands).

Some teachers focused on implementing classroom activities:

Also the dictogloss activity, which obviously we learnt on the LiLAC course, is something we do, I would say, on a weekly basis now (Y6 teacher, North East).

Whereas some focused on the language features that students need to attend to in specific genres:

We were then able to talk about cohesion and about what a letter should look like. So we used this idea that every paragraph in a letter should essentially go red and then yellow and then green. So red opens and introduces it, so it uses a sequence and a conjunctive of some kind and states what the paragraph is going to be about. Yellow then covers my points I want to make. Green then summarises, maybe sets out an expectation, maybe asks a rhetorical question, but then also links and introduces the next paragraph. So then when they actually went through and read their work they had to go red, yellow, and some were going, 'Ah, I haven't linked it' (Y6 teacher, Cambridgeshire).

A number of comments summarised the overall impact of genre-based pedagogy on learning across the curriculum:

I'm the history subject leader here as well and because we've been using it through lots of our history topics, you can see the vast improvements in cross-curricular writing that have been achieved (Y6 teacher, South East).

We are getting those longer, lengthier pieces of writing and more technical language being used, or it looks like an explanation text, but it's in our geography books, rather than just in literacy being taught. So we're seeing an improvement in the cross-curricular writing (School leader, South East).

These comments in particular reveal that the Integrating English training affected different subjects and different genres across the primary curriculum in some schools.

The teaching and learning cycle

The teaching and learning cycle was mentioned in 25 of the 41 interviews, being mentioned five or more times in eight interviews (with an average of 4.3 mentions per interview), and receiving over 5% of discussion time in seven of the interviews. Five of the 19 interviews with teachers that discussed the teaching and learning cycle were interviewed in Y5, so only four interviews with Y6 teachers did not elicit discussion of this topic, making it another salient topic for those trained for Y6.

Discussions around the teaching and learning cycle regularly described the different stages involved:

We're also doing a lot more joint construction as well in pairs and in groups with the children (Y5 teacher, West Midlands).

Such discussions demonstrate how the pedagogic practice had become routine in some classrooms.

The register continuum

The register continuum was discussed in 31 of the 41 interviews. In only two interviews did this topic occupy more than 5% of the discussion, and in five interviews it was mentioned five or more times (averaging 2.9 mentions per interview). Of the ten interviews that failed to mention the register continuum, six were in a leadership position (which is not unexpected, as leaders were less immersed in the programme). Only one Y5 interview did not mention the register continuum, leaving three Y6 teachers that did not consider it relevant enough to mention in their interview.

Structured talk

In 29 of the 41 interviews, the final key principle, structured talk (typically mentioned as talk-as-performance or talk-as-process), was mentioned, taking up more than 5% of the discussion in eight interviews. The topic was not repeated many times, however, with only three interviewees mentioning it five times (and an average of 2.7 mentions). Ten of the interviewees that mentioned talk-as-performance were school leaders of some description and five of the six Y5 teachers mentioned this topic. With 15 of 18 interviews with Y6 teachers mentioning talk-as-performance, it remains a recognisable concept among the majority of the teachers trained in the Integrating English programme for Y6.

Other activity types

Teachers and school leaders in Y6 also talked freely about the types of activities that they had introduced from the Integrating English programme into their classrooms, apart from the main four key principles of the LiLAC programme. These are quantified in Table 19. The mentions are extracted from the interviews in context, as many of these words (such as vocabulary) are used to discuss student progression, school objectives, individual achievement, and other topics. These unrelated mentions were removed from the total. Some of the activities were named slightly differently (see below). The figures are, therefore, open to interpretation and are approximate.

In terms of total mentions, activities that focus on vocabulary (58) were not far behind structured talk (77). While this was not a key principle for the training team, it clearly made an impression on many of the teachers and school leaders, and it was a feature of the LiLAC training programme. Of the individual activities introduced to participants during the training, dictogloss clearly made a disproportionate impact on the teachers' description of their practice, with 46 mentions. This, as noted above, was highlighted as a popular activity.

The next most popular activity was the inference grid (24 mentions) with listening grids or frames (24), Socratic talk or discussion (17), and the barrier task or game (16)—an information gap activity—receiving frequent mentions. Other more general teaching techniques that were frequently mentioned included activities focused on the nominal (or extended noun) group (45), with visual literacy (23) and graphic organisers or visualisers (10) proving popular.

The LiLAC programme attempts to combine an understanding of the role of language in learning across the curriculum with practical classroom activities and techniques. One aspect of the training that emerged from the interviews was the balance between the linguistic concepts behind the LiLAC programme (including register and nominal group structure), the pedagogic framework of the teaching and learning cycle (using the joint deconstruction - joint construction - individual construction sequence), and the teaching techniques, including diamond nine and dictogloss, intended to help implement the more complex concepts.

A slight difference was noted in the discussion of the four main principles between Y5 and Y6 teachers, with Y6 teachers more readily and frequently talking about the four principles. This may be a result of the many discussions reported with Y5 teachers before taking the training in Y6—an advantage not shared by the Y5 teachers. The figures presented here (see Tables 19 and 20 and Figure 3) demonstrate that teachers and school leaders who experienced the LiLAC training were able to refer to some of the more novel and complex concepts, such as genre and the register continuum, as readily as they were able to discuss the impact of teaching techniques, such as dictogloss, on their classrooms.

To summarise, interviews with teachers revealed an enthusiasm for the Integrating English project, with one of the key learning objectives of the training programme, genres and genre-based pedagogy, being recognised by all teachers as a significant aspect of the intervention. However, as with compliance with SoWs and Action Plans, there was wide variation among participants in knowledge and understanding of the other aspects of the training which was also reflected in wide variations in the observations (see below).

Table 19: Mentions of teaching activities and strategies, beyond four main LiLAC principles

Teaching Technique or Strategy	Mentions
Vocabulary focus	58
Dictogloss	46
Nominal / (extended) noun groups	45
Inference grid	24
Visual literacy	23
Listening grids and frames	19
Socratic discussion	17
Barrier task / game	16
Graphic visualiser / graphic organiser	10
Sentence starters	8
Cohesion and pronouns	8
Summary grid	7
Diamond 9	4
Collocations	4
Jigsaw activity	3

Teacher knowledge and practices

During case study visits, in interviews teachers and school leaders were asked about the ideas from the LiLAC training programme and were asked to talk about changes to their teaching practice as a result of the training. In 11 of the 27 interviews, Y6 teachers and school leaders made explicit mention of the skills and knowledge they had gained from the training programme.

Teachers were split between those, in the minority, that felt that there was nothing new in the LiLAC programme:

I think it's just nice refreshing. Some of the strategies you maybe did know, because I've been teaching longer (Y6 teacher, South East).

Then there were those that felt that LiLAC offered a highly appreciated review of key pedagogical principles and practices:

but what I think this has done, it's just made me and my colleagues more aware of the purpose of what we're doing. It's having a purpose of why you're doing this activity, how is it going to help the students (Y6 teacher, North East).

So that teaching and learning cycle I think is a lot clearer to me than it was before the training (Y6 teacher, Leicester / Peterborough).

Some were impressed by the impact of the training on their practice:

I think it's had a really big impact on my own teaching. It's definitely been the best bit of training I've been on since I started. This is my third year as a teacher now and I've been on some very good training and I think this is probably the best training I've been on still. It was really great to be able to be exposed to so much good theory and useable practice in the training (Y5 teacher, North East).

In these comments, and drawing on the quantitative summary of the interview data (Implementation of Training), all of the main elements of the training for the Integrating English programme (genre-based teaching, teaching and learning cycle, register continuum and talk-as-performance) contributed to most teachers' knowledge and understanding, with genre-based teaching remaining salient for all teachers. For some who claimed to be aware of these ideas, the training provided a review, reintroduction, or initial training in these areas. A small number of teachers and leaders focused specifically on the contribution the training made to teaching EAL students, whether through specific strategies or through awareness of the challenges that EAL students face. Others made frequent mention of specific teaching techniques and routines (see Table 20) that they found to be of value.

Teachers and leaders were also encouraged to consider any changes in belief or attitude they had experienced as a result of the Integrating English project. A number of comments focused on new perspectives when teaching EAL students. Some of these related to removing a 'deficit' view of EAL students:

I think it opened our eyes to the difficulties that EAL students can face when it's not the language that they speak at home every single day. So they are transitioning from being at home and some of them speaking Arabic fluently, to then obviously being in school and speaking English as fluently as they can. I think it opened our eyes as teachers to taking different approaches and giving them that time to talk and that time to embed and that time to feel secure in trying out the language, before it being in a formal way where they maybe wouldn't use it because they're scared to get it wrong. The building up, the joint construction, we do a lot of that now in the building up, whereas a lot of the time you start at the top (Y6 teacher, West Midlands).

Others related to seeing EAL students as having something to offer:

but what's really been brought home to me is that EAL students come with already a lot of knowledge. We sometimes think they don't speak English so we need to put them in the lower group and that's not the case. We've got to remember that they've got a lot of knowledge themselves in their own language and we need to use that, not shy away from it. That really is what's been brought home to me, and I'm more aware of that now (Y6 teacher, North East).

Further comments revealed a reorientation to the curriculum and to EAL students in general:

Sometimes we think we have to give them simplified language and we don't. Actually we don't set high enough standards for them. I've seen, by setting that high standard, they will really acquire that formal language to start with and they are actually ahead of some of the other children at the end, because they have now just learnt the formal language and they know how to use it in context. It might take them slightly longer sometimes, but when they've got it, they can use it across the board (Y6 teacher, South East).

A number of comments also made it clear that the gap between EAL students and other disadvantaged groups was not that great, and that many of the solutions that Integrating English provided worked for a range of students:

What I loved about the project is yes, it was essentially for children who are acquiring English, but, as I said, the benefits were there for all pupils to be had. It wasn't exclusive to them, so actually lots of children benefitted from this project, not just the target audience (School Leader, South East).

This suggests that for some of those involved in the programme, there were numerous examples of changes in teaching knowledge and practices, and attitudes towards students from different groups, including EAL students, changed considerably as a result of the training.

Pupil outcomes

Teachers and school leaders were asked to report on changes that they had observed in their students as a consequence of the Integrating English intervention. Responses ranged from suggesting that no clear changes had been noticed, most frequently in the area of behaviour, to reports of students claiming they enjoyed working within the teaching and learning approach promoted by the Integrating English programme.

Teachers and school leaders in intervention case study schools were asked directly if they had noticed an impact on student behaviour or on their confidence and engagement (see the logic model in Appendix C). Very few comments were made about student behaviour; teachers typically responded that they experience few behavioural problems at their school. One teacher described how a particular student could prevent lesson objectives from being achieved as a consequence of habits at home, while one headteacher suggested that the combination of Integrating English and the Success for All programme had achieved significant behavioural changes across the school.

In about 50% of the case study schools, teachers and school leaders made clear comments on improvements in the confidence of their students as a result of the pedagogic strategies introduced by the Integrating English programme. Some teachers highlighted both speaking and listening activities where they could see their students showing greater confidence as a result of how Integrating English strategies structure these activities. A number of individual students were highlighted by other teachers and one of the school leaders offered this summary:

I always think in terms of outcomes, so forgive me, in terms of academic outcomes, but there are huge outcomes in terms of confidence and happiness (Y6 school leader, South East).

About 50% of the case study schools offered comments on engagement, too. Many of these teachers described how the teaching strategies they were introduced to in the Integrating English training require all of the students to participate. Because there is no option to allow the stronger students to take over, all students have become more engaged and increased their participation. The EAL students were mentioned in particular by some of the teachers for having increased engagement as a result of tasks that scaffolded the required language:

The children who had a real paucity in language—there weren't that many in my class, but there were a few in the other class—I really noticed they were more engaged because they could access the learning more (Y6 school leader, South East).

A teacher at one of the schools described how some Integrating English activities challenged the students to concentrate or remain focused for an extended period and that this could result in a lack of engagement, suggesting that the students had too short an attention span. Another teacher remarked that the approach promoted in Integrating English had already narrowed the gap between her students:

it was an EAL child and a higher ability child having that discussion about where it [a word] should go, so that shows me actually they're both thinking now about when it's appropriate to use and how we use it (Y6 teacher, South East).

The comments included here are, of course, only observations made by the teachers and school leaders in the case study schools of their students' behaviour, confidence, and engagement. It is possible that these developments may have taken place due to natural maturational and learning cycles, a point made by one of the school leaders, or that the project interview had focused their attention on an area that would not normally receive as much interest.

School-wide changes

The interviewees were asked to consider the effect of the Integrating English project on the school generally—an effect predicted by the logic model for the intervention (Appendix C)—and to comment on how practice in the school may have been enhanced due to the project. As discussed previously, a number of comments revealed how easily the Integrating English training integrated with other literacy interventions. A few interviewees talked about how some ideas, such as the inference grid or the language line (or register continuum), had already been rolled out across the school, helping students with their transition from one teacher to another. Some schools made mention of including all teaching staff, particularly teaching assistants, in enhancing school practice, while others pinpointed targets for further improvement, such as spreading Integrating English training to more areas of the curriculum.

Interviewees talked openly about how they had passed on aspects of the Integrating English training to other teachers in the same school, often resulting in the adoption across the school of certain practices. For instance, one school mentioned that the register continuum had been introduced to all classes in KS2. When asked about sharing practice around the school, most schools responded that they had already started sharing the ideas from the Integrating English training with their colleagues, particularly as Y5 teachers handed over to Y6. A teacher at the one school that was generally negative about the whole programme commented that they would not do so, however, this view was very much in the minority. Ideas and techniques that had already been shared within different schools included:

- barrier task;
- diamond 9;
- expanded noun phrase;
- formal and informal language;
- guided reading;
- inference grids;
- Socratic talk; and
- summary grid.

It is interesting to note from this list, however, that in-school CPD was not seen by participants to provide a suitable platform for the sharing of the four main principles of LiLAC training. While there was a lot of enthusiasm for sharing practice, with some schools describing their systems for ensuring in-school CPD, teachers and leaders from other schools lamented the lack of time for, and the lower priority given to, CPD.

Observations

Visits to case study schools produced 31 observations of different classrooms with different teachers across the five regions in Years 5 and 6. In all cases, observers were invited to attend the classes identified by the school as using principles, ideas, or techniques from the Integrating English training programme. The aim of the observation tool (Appendix G) was to gather evidence of implementation of the programme, focusing on:

- explicit attention to language during the class, and whether it would benefit EAL students in particular or all students more;
- the use of genre to describe or explain language variation;
- the implementation of genre pedagogy or the teaching-learning cycle;
- student engagement; and
- teacher confidence in discussing language.

This section summarises the data from these observations.

Focus on language and genre

In lessons that observers attended, all classes exhibited some kind of task, activity or discussion that focused on language. The most common focus was on the register continuum and the structure of the nominal group or how to make a noun more complex. Definitions of vocabulary items were explained or elicited in a range of classrooms and across different stages of the lesson. Combining register and vocabulary, there was evidence of distinguishing and exploring technical vocabulary in some classes. Other more grammatical examples of a focus on language included discussions or activities around verb tense, sentence structure, pronouns, and question formation, but these were typically isolated instances.

In a small proportion of classes, students worked explicitly with genres. Genres that were named in class included crime fiction, diaries, non-chronological reports, explanations, and discussions, and in most cases classes identified features associated with the genre. For instance, in one lesson pupils distinguished features that could or could not be expected in a diary, and another class identified relevant sequencing words in a report. Technical vocabulary and the register continuum also featured in some of the genre-focused activities.

Observers noticed that in some classes certain language-based activities seemed to benefit EAL students more. These included predicting vocabulary before a listening exercise, identifying tense and pronouns, and focusing on definitions. In some cases, these same language-focused activities were considered of value to all students, while paying attention to the structure of nominal groups was considered valuable for all students.

Genre pedagogy and the teaching-learning cycle

As well as the linguistic features of a genre, the Integrating English programme expects teachers to expose students to the social role of texts in context during the deconstruction stage of the teaching and learning cycle (see Figure 1). However, none of the case study observations witnessed this stage in the classroom. Evidence of individual construction was supplied both with 'cold writing tasks' (set before students learn about a topic and how to write about it) and final products from a sequence of lessons using Integrating English techniques and strategies. There were, however, a number of observations of joint construction writing, where students work together in groups or the class works with the teacher to produce a text.

Observers frequently witnessed students being given time to plan for writing, as well as planning for presenting, also known as 'talk-as-performance'. In these classes, students were evidently familiar with this type of preparation, and teachers provided various tools, such as spider-grams for brainstorming and fishbone diagrams to aid planning. The time given for planning was judged to be of value to all students alike. There was, then, evidence that some, but not all, stages of the teaching-learning cycle were common in these classes.

Students and teachers working with language

Observers tried to gather evidence of whether teachers and students felt comfortable when working with language, looking particularly at teacher confidence and student engagement. In a few cases, there was evidence that the teacher was not confident about providing explanations about genre or grammar, but in at least ten cases the teacher was observed as being confident in their description of language, including talking about nominal groups, sentence structure, and vocabulary.

In general, observers found that students readily engaged with the activities in class and completed tasks with interest and enthusiasm. In one or two cases, individual students were identified as not being on-task and one task was identified as being too poorly explained for students to complete it, even though they had previously worked well on all other tasks. Students were observed employing the metalanguage used in the LiLAC programme and using inferencing skills in an 'inference frame'

activity—another technique introduced as part of the Integrating English programme. It should be noted, however, that in the majority of schools, teachers reported few problems with engagement or discipline in class and where issues arose this was often attributed to one or two individuals that were inattentive in all classes.

Pre- and post-intervention survey

The evaluation team intended to measure differences in attitude between teachers in the control and intervention groups, and between intervention group teachers at the start and the finish of the project (see Appendix F). Unfortunately, the team did not receive a sufficient number of post-intervention responses to make any claims that would bear statistical scrutiny.

A pre-intervention survey of teachers was sent only to Y5 teachers in participating schools (intervention and control) prior to randomisation. The questionnaire covered teaching practice and CPD training experience around literacy and EAL pupils and was distributed electronically. A total of 117 valid responses were received. The balance between intervention (n = 58) and control (n = 59) teachers was good. The teacher-level response rate cannot be calculated accurately as the survey was disseminated through the project lead in each school. Responses were received from 67 schools, giving a school-level response rate of 74%.

A survey of Y6 teachers was conducted in the summer of 2016, after they had been identified, yielding 47 responses. This was undertaken at a time when Y5 teachers had already attended LiLAC training and commenced teaching as part of the intervention programme and could not therefore be treated as pre-intervention data despite the respondents not having yet had direct involvement with the intervention through training. For further details on the delivery schedule, please see timeline earlier in the report.

The post-intervention survey was sent to schools following the completion of primary outcome testing. It was agreed that the timing of this was sensible given intense efforts required to coordinate primary outcome testing in schools. One consequence of this decision is that the survey was distributed close to the end of term. Only 19 valid responses were received. Of these, only ten teachers had also completed the baseline survey—three from control schools and seven from intervention schools. It is unclear why the response rate was so poor. The timing, so late in the term, is perhaps one explanation; another reason could be that teachers felt as though they had had enough of the evaluation demands and did not respond as a result.

While a respectable response rate was achieved in the pre-intervention survey, reporting the results here could reveal interesting findings about the situation in participating schools, but would tell us nothing about the effect of the intervention. The ten respondents who participated in both the pre- and post-intervention survey are likely to be atypical in terms of engagement with the programme. Possibilities for analysis are severely restricted by the small sample size and any comparison of pre/post intervention differences between the intervention and control groups would be rendered meaningless by bias.

However, the baseline survey results indicate that across a total of seven survey items relating to teacher knowledge, confidence, and practice, the responses of teachers from intervention and control schools were well balanced. On items where differences were observed, the chi-square statistic for the cross-tabulation was not statistically significant. We therefore have no evidence that business-as-usual would have been significantly different between intervention and control schools. The tables are reported in Appendix I.

Formative findings

The LiLAC training programme is ‘imported’ from a different educational context, Australia. Some of the training materials were identifiably Australian because of the accents, classrooms, and clothing of video participants. While the trainers worked hard to make the training materials more relevant to the U.K. educational context, it is rare that any teaching methodology can be directly transferred across cultures easily (Holliday, 1995). Various terms were adapted to the U.K. context, and recognition was repeatedly made of the increase in EAL students in U.K. schools and of the pressure of league tables, Ofsted inspections, and measures of improvement including SATs for Y6. However, frequent complaints were made in interviews about the terminology used in the training not matching up with the standards used in national examinations for Y6 students, resulting in teachers using U.K. terms for LiLAC concepts. These may, at times, be simple replacement terms for pedagogic effect: it is far simpler for Y5 pupils to say *extended noun group* than *nominalisation* for instance. At other times, however, the functional motivation behind the labels may be lost in this apparent simplification; replacing register with ‘language line’ may prevent a discussion of what variables change a register. A review by the training team into possible alternative terms, and perhaps a glossary identifying where the LiLAC terms differ from those in Grammar, Punctuation and Spelling (GPS) tests, may help everyone understand how and why there are different terms for seemingly similar items.

While all teachers interviewed at case study schools discussed the principle of genre pedagogy, and it was recognised as a central element by the vast majority of school leaders, there is not a lot of evidence to show that these principles were being enacted. For instance, none of the observers witnessed text deconstruction, or joint or individual reconstruction—key stages in the genre pedagogy cycle. Further, these stages were rarely itemised on schemes of work. It is possible that as stages in a lesson, teachers may have felt no need to make the step explicit, but it is not unreasonable to assume that at least in earlier lessons, as teachers grappled with introducing Integrating English into their curriculum, they would need to be explicit about how they would implement the many different ideas. The trainers expected teachers to use at least three Integrating English ideas a week, so the pressure would be on teachers to write down all instances of Integrating English pedagogy. It may be possible for LiLAC trainers to put more emphasis on the four main principles of the training, or it could also be the case that U.K. primary teachers do not have the required grammatical knowledge from their pre-service and wider in-service training to identify key textual features in different genres. To increase the likelihood of teachers implementing the key principles of Integrating English, more than the teaching techniques, into their classrooms, Integrating English trainers may need to be more explicit about challenging the pedagogic norms of U.K. classrooms in order to be more explicit about how genre pedagogy is different. Another solution may be to focus training on the features of discipline-specific genres so that teachers will be better equipped to apply genre pedagogy and functional linguistic principles to new texts encountered during the school year.

Regardless of the results of the randomised controlled trial, all teachers and school leaders—with one exception—were very complimentary about the training programme, the trainers, and the effects that they had perceived on their teaching, the school in general, and their pupils. While the response rate to survey questions was too low to examine wider perspectives, comments provided by case study schools about the training were overwhelmingly positive, with recognisable changes in engagement and confidence, particularly among EAL students. The process evaluation suggests, therefore, that teachers perceive the programme as being effective in more than just test scores. Further, as we note in the concluding section below, the use of schemes of work to act as a bridge from training to practice is an interesting model for future studies; in this case they were variably produced but the positive examples indicate the potential of the development of curriculum materials such as schemes of work as part of the implementation process in future intervention designs.

Control group activity

Schools in the control group were considered to be carrying on 'as usual'. Whilst we have no evidence that the LiLAC/Integrating English approach in particular is used widely in England, including among the control group, data from the case studies reported above indicates that a number of schools felt the approaches they already used (for example Success for All, Talk for Writing) aimed for outcomes similar to the Integrating English programme and others felt their current approaches were consistent with Integrating English. Given the random allocation of schools to intervention and control groups, it is plausible that a similar number of control schools had similar approaches in place, given that findings from the pre-intervention baseline survey indicated that relevant teacher confidence, knowledge, and practice was well balanced between the intervention and control groups.

Conclusion

Key conclusions

1. There is no evidence that Integrating English improved pupils' KS2 writing outcomes. This result has a moderate to high security rating.
2. There is no evidence that Integrating English had an impact on the KS2 writing outcomes of pupils receiving free school meals. These results have lower security than the overall findings because of the smaller number of pupils.
3. There is no evidence that Integrating English improved EAL pupils' KS2 writing outcomes. Although this was measured through a large subgroup analysis, these results have lower security than the overall findings because of the smaller number of pupils.
4. The process evaluation indicates that, although teachers responded positively to the training, the CPD model may not have been effective in creating the desired teacher practice change. A simpler model may be more effective.

Interpretation

The IPE evaluation is always important in trial-based studies, but it has particular importance in relation to studies such as this for which there is no evidence of improvement. Coldwell and Maxwell (2018) draw on a well-established approach from the policy evaluation field—they discuss Lipsey and colleagues' (1985) 'three forms of failure'—to distinguish three sets of reasons, which are not mutually exclusive, that might apply in cases where there appears to be no impact:

- theory deficiency: the causal theory underlying the intervention is faulty;
- implementation deficiency: the implementation of the intervention is such that the intervention is not applied in such a way that it can lead to positive outcomes; and
- methodological deficiency: the evaluation design and application is at fault—it is measuring the wrong things, or is so limited that it cannot detect impact.

It is helpful at this stage to return to the evidence-informed logic model. As noted earlier, the model for the intervention has two distinct causal processes. First, a professional development programme consisting of a set of external events—with associated development of scheme of work and action plan—was expected to lead to changes in teaching practices via changed knowledge exemplified in four principles of the programme. Second, these changes in teaching practices were expected to lead to improved literacy outcomes for children, and to embedding in wider school processes.

It is clear from the IPE that the first of these processes did not lead consistently to the changes in practice that were hoped to translate, in turn, to impacts on pupils and wider school processes. Whilst training was experienced positively, the crucial move from translating the learning into practices via action planning and scheme of work development was variable. There was evidence that part of the issue here was one of implementation deficiency: a complex programme reliant on engagement in a series of elements has multiple potential points of failure including attendance at training, production of an appropriate action plan, production of an appropriate scheme of work and engagement in Fronter. There was evidence of a lack of consistency at each of these points from the case study component of the IPE study. One indicator of this was that there was poor fidelity using the range of measures agreed between the evaluators and delivery team as only 12 of the 45 intervention schools were judged to have been fully compliant with the programme (although even in these 12 schools no evidence of impact was found). It could also be argued that a weakness in the programme was the lack of any in-built monitoring of in-class implementation, such as through action research cycles, or that the 'between-sessions' tasks, where teachers were invited to practice ideas from the Integrating English training, were inadequately supported or were insufficient to expect confident implementation of the wider programme.

Future interventions based on LiLAC or Integrating English should consider how to further support in-class implementation.

Beyond this, however, there was evidence from the IPE of a wider problem with the underlying change mechanism inherent in the professional development element. There was a range of evidence that teachers addressed the underlying programme principles variably, focusing much more on genre pedagogy than the others (Figure 3), with variable evidence that the schemes of work, in particular, represented understanding of the programme principles. The result of these issues was that there was variable evidence from discussion and observed practice in case study schools in relation to practice changes that focused on key relevant processes such as deconstruction and reconstruction of texts. Overall there was evidence that some, but not all, stages of the teaching-learning cycle were common in observed classes: teachers mainly responded to strategies, hints and tips rather than introducing a new pedagogic cycle, and there was little evidence that teachers had the linguistic knowledge to deconstruct school genres in order to expose these linguistic features to pupils.

Given the highly variable evidence of implementation of practice in line with LiLAC principles, we would expect variable outcomes for schools and pupils. Broadly this means that, whilst the prior Australian research evidence summarised in the background section indicates that the underlying theory in relation to the potential usefulness of genre pedagogy approach has a clear theoretical basis and some empirical corroboration, we have much less evidence in relation to this second causal process from the current study. Since we did not see the changes in teachers' practices, we have limited evidenced on whether these practices would have led to positive pupil change had they been seen.

This indicates that the model of professional development was not effective in creating the anticipated teacher practice changes, and alternatives should perhaps be used to improve this process. This might involve clearer subject targeting: a LiLAC-influenced approach in relation to a specific subject area (such as science, mathematics, history, geography) may provide more opportunity for the approach to lead to more consistent practices. Further, a simpler model which involved an implementation process with fewer steps would reduce some of the potential for problems. Thought should also be given to the inclusion—as here—of two sets of teachers, one in Y5 and one in Y6, each of which going through this process. Whilst this had the potential to reinforce learning and embed it in the school, it also increased the possibility for issues of fidelity and complications in the implementation. Finally, the dosage should be considered. Considering the variety of objectives for the Integrating English training programme, and the lack of a knowledge of functional grammar in most U.K. primary teachers, a 50-hour CPD course may not be enough to affect sufficient long-term change in teachers' linguistic knowledge. Implementation in a single term in one subject area in two separate years may also be an insufficient length of time to produce significant changes in pupils' performance in subject-based literacy.

The role of the scheme of work deserves special mention here. Whilst schemes of work were variably produced, they appear to have the potential to act as an appropriate method to smooth the enactment of professional development into practice by incorporating new approaches into existing school practices and broader curricular aims in a standardised framework that makes pedagogic objectives and priorities transparent to a range of stakeholders. The variability of enactment, however, suggests that more work is needed on developing this part of the programme.

Limitations

The impact evaluation adopted a robust Clustered RCT design and evidence for methodological failure is limited to two aspects: the lack of detail on the 'business as usual' control condition and missing data. The data on what constitutes 'business as usual' was dependent on a post-treatment survey which was returned by insufficient respondents in both control and treatment groups, although findings from the pre-intervention teacher survey indicated balance between intervention and control schools in teacher confidence, knowledge, and practice. Factors such as timing may have contributed to the low response rate to the post-intervention survey, yet future data collection for the context of the control group needs

to be more robust and depend on more than a MoU. While details of 'business as usual' may be scarce, certain assumptions can be made fairly safely. One of these is that all schools taking standardised KS2 tests will implement strategies and approaches to maximise scores. The impact evaluation shows us that the Integrating English approach produced no gain in KS2 GPS and KS2 Reading scores over these other methods.

Missing data for the primary KS2 writing outcome was predominantly explained by intervention schools dropping out of the intervention and therefore not participating in the outcome testing. We did not find that these missing schools resulted in a notable imbalance between the intervention and control school samples. Within schools that did participate in the outcome testing, we found no evidence that the pattern of missing data in intervention schools was different from that seen within the control schools. Finally, the secondary outcomes (KS2 GPS and Reading) were drawn from the NPD and so did not suffer the same extent of missing data (4%). For both these secondary outcomes, we also conclude no evidence that Integrating English resulted in pupil gains in attainment.

A few comments in interviews, and comments from the training delivery team, questioned the relevance of a general writing test to measure students' progress against the control groups since the strength of the Integrating English programme, in writing, was not necessarily in writing typical of English and 'literacy' contexts but across the curriculum in different subjects. The variety of curriculum subjects that Integrating English was applied to in different schools would make it next to impossible to find a suitable writing test for all.

Moreover, since the Department for Education decided with little or no consultation to cancel the statutory demand for all schools to take what was the earlier SPaG (Spelling, Punctuation and Grammar) test, the replacement optional English GPS (Grammar, Punctuation and Spelling) test essentially became a lower-stakes examination; it is likely that a similar number of pupils across both groups did not take the writing test seriously. However, despite these concerns it is worth reiterating that pupil attrition for the KS2 GPS test was as low as for KS2 Reading, which remains compulsory. It should also be noted that no effect on pupil progress was found on the Writing assessment used as the primary outcome, KS2 GPS or KS2 Reading.

Trainers and participants were concerned that evaluation of the programme may need to consider a more suitable tool for primary and secondary measures of change since it can be safely assumed that teachers in both control and intervention groups aimed for the best results in the high-stakes SATs tests of general English reading and writing. The impact evaluation results suggest that the Integrating English programme was no more effective in raising these scores than any other approach in the control schools. However, what was not measured in these tests were students' skills in subject-specific literacy, which was the focus of the intervention. The key problem (for the purposes of the trial) was that a range of subjects were used in schools. A validated tool that accurately measured responses in a comparable way in relation to the specific subjects in the range of subjects addressed in schools was not available (and it is doubtful such a tool could be created, not, at least, without a huge amount of developmental work). Without such a tool it is difficult to measure literacy development across subjects.

Future research and publications

As indicated above, future research studies may profitably examine whether alternative outcome measures could be used, as indicated in the limitations section above. Furthermore a simpler, more intensive teacher development programme that aims to introduce changes to teacher practices in relation to genre pedagogy with a focus on a single subject area (for example science) may be worthwhile.

Finally, as indicated earlier, there are indications in this study that more specific research examining the potential for the use of schemes of work (and possibly other, related, curriculum planning approaches) as a way of bridging CPD and teacher practices in school contexts would be worthwhile.

References

- Bernstein, B. (1975) *Class Codes and Control, Vol. III: Towards a Theory of Educational Transmissions*, London: Routledge.
- Bernstein, B. (1981) 'Codes, modalities, and the process of cultural reproduction: A model', *Language in Society*, 10 (3), pp. 327–363.
- Bernstein, B. (1990) *Class, codes and control, Vol. IV: The structuring of pedagogic discourse*, London: Routledge.
- Bloom, H. (1995) 'Minimum Detectable Effect Sizes: A simple way to report the statistical power of experimental designs', *Evaluation Review*, 19 (5), pp. 547–556.
- Bloom, H., Richburg-Hayes, L. and Rebeck Black, A. (2007) 'Using covariates to improve precision for studies that randomise schools to evaluate educational interventions', *Education Evaluation and Policy Analysis*, 29 (1), pp. 30–59.
- Christie, F. and Derewianka, B. (2008) *School discourse*, London: Continuum.
- Coffin, C. (2006) *Historical discourse: The language of time, cause and evaluation*, London: Continuum.
- Coffin, C., Acevedo, C. and Löfstedt, A. (2013) 'Teacher Learning for European Literacy Education (Tel4ELE)', Final Report for Comenius Multilateral Project.
<<https://www.readingtolearn.com.au/wp-content/uploads/2016/01/Final-Report-Teacher-Learning.pdf>>
- Coffin, C. and Donohue, J. (2014) 'A language as social semiotic-based approach to teaching and learning in higher education', *Language Learning Monograph Series, Language Learning*, 64 (s1).
- Coldwell, M. and Maxwell, B. (2018) 'Using evidence-informed logic models to bridge methods in educational evaluation', *Review of Education*, 6 (3), pp. 267–300.
- Custance, B., Dare, B. and Polias, J. (2012) *Teaching ESL students in mainstream classrooms: Language in learning across the curriculum*, Hindmarsh: DECD.
- Demie, F. and Strand, S. (2006) 'English language acquisition and educational attainment at the end of secondary school', *Educational Studies*, 32 (2), pp. 215–231.
- Derewianka, B. (2012) 'Knowledge about Language in the Australian Curriculum: English', *Australian Journal of Language and Literacy*, 35 (1), pp. 127–146.
- Halliday, M. (1978) *Language as social semiotic: The social interpretation of language and meaning*, London: Edward Arnold.
- Halliday, M. A. K. (1993) 'Towards a language-based theory of learning', *Linguistics and Education*, 5, pp. 93–116.
- Halliday, M. A. K. and Hasan, R. (1989) 'Language, context, and text: aspects of language in a social-semiotic perspective', Victoria: Deakin University.
- Halliday, M. and Matthiessen, C. (2014) *Halliday's introduction to functional grammar*, London: Routledge.
- Hardman, F. and Hardman, J. (2017) 'Observing and recording classroom processes', in D. Wyse, N. Selwyn, E. Smith and L. Suter (eds), *The BERA/Sage handbook of educational research*, Los Angeles: SAGE (ch. 27).
- Holliday, A. (1995) *Appropriate methodology and social context*, Cambridge: Cambridge University Press.

- Jo, B., Asparouhov, T., Muthén, B. O., Jalongo, N. S. and Brown, C. H. (2008) 'Cluster randomized trials with treatment noncompliance', *Psychological methods*, 13 (1), pp. 1–18.
- Lipsey, M. W., Crosse, S., Dunkle, J., Pollard, J. and Stobart, G. (1985) 'Evaluation: The state of the art and the sorry state of the science', *New Directions for Evaluation*, 27, pp. 7–28.
- Locke, T. (ed.) (2010) *Beyond the Grammar Wars*, London: Routledge.
- Macken-Horarik, M., Love, K., Sandiford, C. and Unsworth, L. (2018) *Functional Grammaticals*, London: Routledge.
- Macken-Horarik, M., Sandiford, C., Love, K. and Unsworth, L. (2015) 'New ways of working "with grammar in mind" in School English: Insights from systemic functional grammatics', *Linguistics and Education*, 31, pp. 145–158.
- Martin, J. R. (1985) *Factual Writing: Exploring and Challenging Social Reality*, Oxford: Oxford University Press.
- Matthiessen, C. M. I. M. (2006) 'Educating for advanced foreign language capacities: exploring the meaning-making resources of languages systemic-functionally', in H. Byrnes (ed.), *Advanced Language Learning*, London: Continuum.
- Maxwell, B., Clague, L., Byrne, E., Culliney, M., Coldwell, M., Hobson, A. and Glentworth, A. (2008) 'Retain: CPD for Early Career Teachers of KS1', London: EEF.
<https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Retain.pdf>
- Myhill, D. and Watson, A. (2014) 'The role of grammar in the writing curriculum: A review of the literature', *Child Language Teaching and Therapy*, 30 (1), pp. 41–62.
- Nebauer, A. and Sungaila, H. (1980) 'The implementation of the disadvantaged schools program in an Australian country region: an evaluation', *Journal of Education Administration*, 27 (1), pp. 165–167.
- O'Halloran, K. L. (2004) *Mathematical discourse: Language, symbolism and visual images*, London: Bloomsbury.
- Pomagalska, D. (2019) 'Evaluation of the whole-school implementation of a language-based pedagogy at Southern Secondary College'. <https://ca4p02htfsc1lr0aengrwwdh-wpengine.netdna-ssl.com/wp-content/uploads/2019/04/Report-LfL-implementation-at-Southern-Secondary-College-Lexis-Education-2019.pdf>
- Prince, P. (2013) 'Listening, remembering, writing: Exploring the dictogloss task', *Language Teaching Research*, 17 (4), pp. 486–500.
- Rose, D. and Acevedo, C. (2006) 'Closing the Gap and Accelerating Learning in the Middle Years of Schooling', *Literacy Learning: The Middle Years*, 14, (2), pp. 32–45.
- Rose, D. and Martin, J. (2012) *Learning to write, Reading to learn: Genre, Knowledge and Pedagogy of the Sydney School*, Sheffield: Equinox.
- Saldana, J. (2011) *Fundamentals of Qualitative Research*, New York, NY: Oxford University Press.
- Schochet and Chaing (2011) 'Estimation and identification of the complier average causal effect parameter in education RCTs', *Journal of Educational and Behavioral Statistics*, 36 (3), pp. 307–345.
- Smith, K., and Davies, J. (2010) 'Qualitative data analysis', in L. Dahlberg and C. McCaig (eds), *Practical Research and Evaluation*, London: Sage.
- Strand, S., Malmberg, L. and Hall, J. (2015) 'English as an Additional Language (EAL) and educational achievement in England: An analysis of the National Pupil Database.', London:

EEF.

https://educationendowmentfoundation.org.uk/public/files/Publications/EAL_and_educational_achievement_Prof_S_Strand.pdf

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A. and Raudenbush, S. (2011) 'Optimal Design Plus Empirical Evidence: Documentation for the "Optimal Design" Software'.
<http://wtgrantfoundation.org/library/uploads/2015/11/OD-Documentation-V3.pdf>

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

OUTCOME: KS2 Writing (Raw scores using previous year's papers)

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

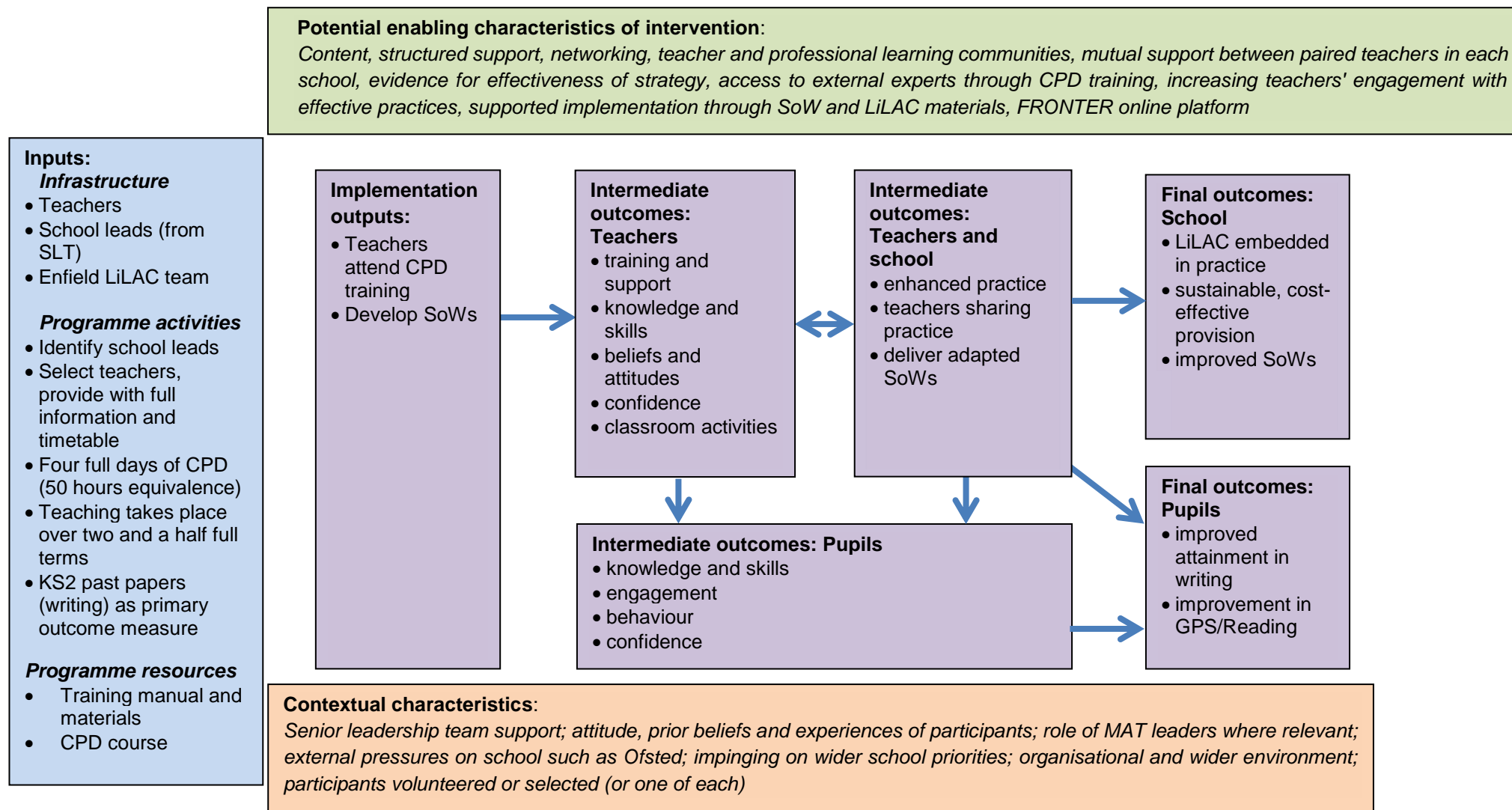
Threats to validity	Threat to internal validity?	Comments
Threat 1: Confounding	Low	- Baseline imbalance was low (+0.04) and controlled for analytically in the model which did not affect the results. - Appropriate description of the randomisation procedure, even if analyst was not blinded at analysis
Threat 2: Concurrent Interventions	Moderate	Data on the case studies reported that some control schools felt that approaches aimed at similar outcomes were already in use (e.g. 'Success for All' and 'Talk for Writing')
Threat 3: Experimental effects and contamination	No information available	The data on what constitutes 'business as usual' and the exploration of changes over the course of the study was dependent on a post-treatment survey which was returned by insufficient respondents in both control and treatment groups to be deemed usable.
Threat 4: Implementation fidelity	Moderate	-The IPE indicates that the teacher development programme did not consistently lead to teacher practice changes that were theorised to produce the pupil and school changes sought. -The complex programme was reliant on engagement with multiple elements including attendance to training, action plan, schemes of work, among others. However, these were implemented with some variability

Threat 5: Missing Data	Low	No evidence that the pattern of missing data in intervention schools was different from that seen within the control schools. However, there is some evidence that missing information was associated to EAL and FSM pupils.
Threat 6: Measurement of Outcomes	Low	Reliable, high-stakes test was used. Nevertheless, the delivery team questioned the relevance of a general writing test to measure students' progress against the control groups, since the strength of the LiLAC programme, in writing, was not necessarily in writing typical of English and 'literacy' contexts but across the curriculum in different subjects.
Threat 7: Selective reporting	Low	This trial was registered and all analyses were conducted as specified in the Protocol and SAP.

- **Initial padlock score:** [3] Padlocks - Randomised design, 0.18 MDES, 24% pupil attrition (primary outcomes).
- **Reason for adjustment for threats to validity:** [0] Padlocks – The training programme showed a high level of completion. However, the measures of fidelity in the process evaluation suggest that there were only a minority of schools that were fully compliant with the demands of the programme. Information from some case studies suggested that similar approaches were implemented in schools. However, none of these were deemed severe enough to reduce the security rating of this study.
- **Final padlock score:** initial score adjusted for threats to validity = [3] Padlocks

Appendix C: Logic model

Theory of change-based Logic model for Integrating English, July 2016



Appendix D: Distribution of baseline and outcome measures

The primary outcome was Writing, using the from the KS2 test paper 2012 academic year.

Table D1 and Figure D2 summarise the distribution of the raw primary outcome variable. The distribution of test scores was normal for both the control and intervention groups.

Table D1: Summary of the distribution of KS2 writing attainment

Allocation	Missing n	Missing %	Outcome N	Mean	SD
Control	489	21.5%	1790	15.79	6.187
Intervention	636	25.9%	1817	15.72	5.916
Total	1125	23.8%	3607	15.75	6.051

Figure D1: Distribution of GL PTE L9 score (baseline measure)

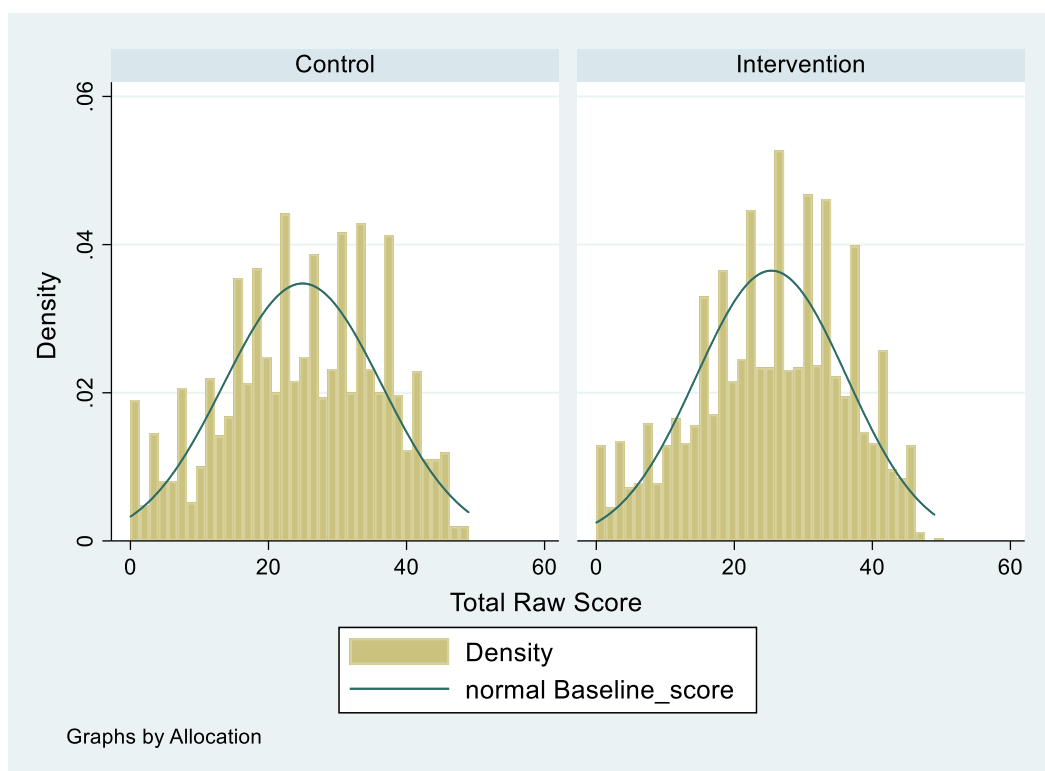


Figure D2: Distribution of KS2 writing attainment (primary outcome)

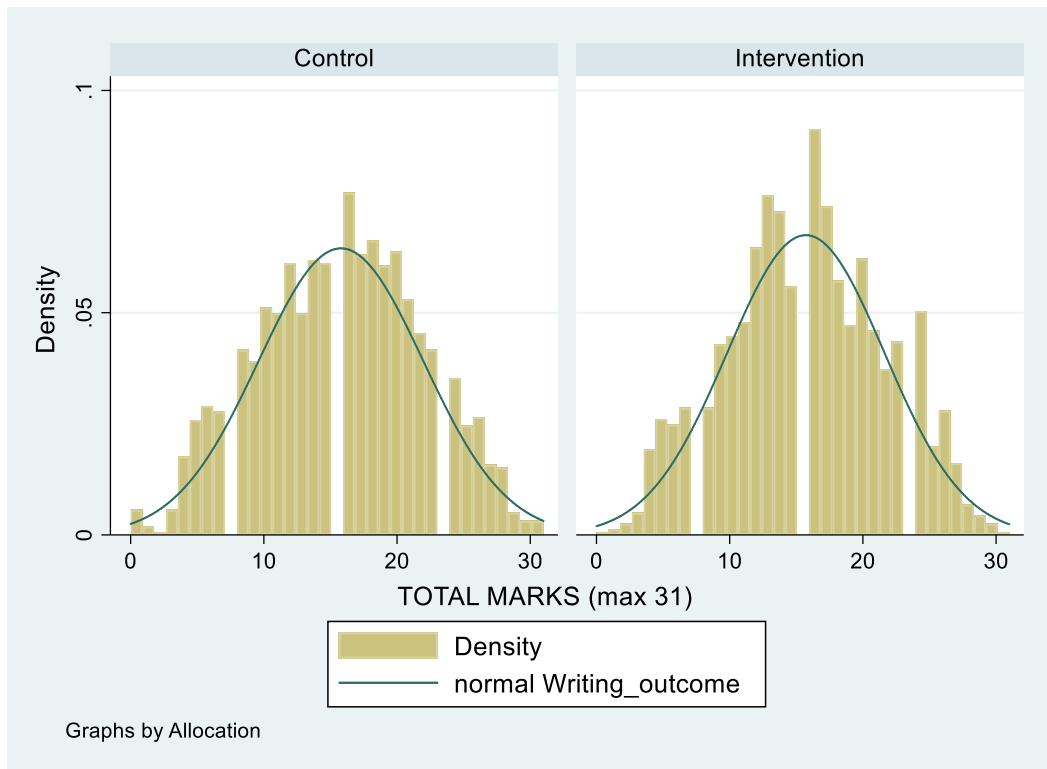


Figure D3: Distribution of KS2 Reading (secondary outcome)

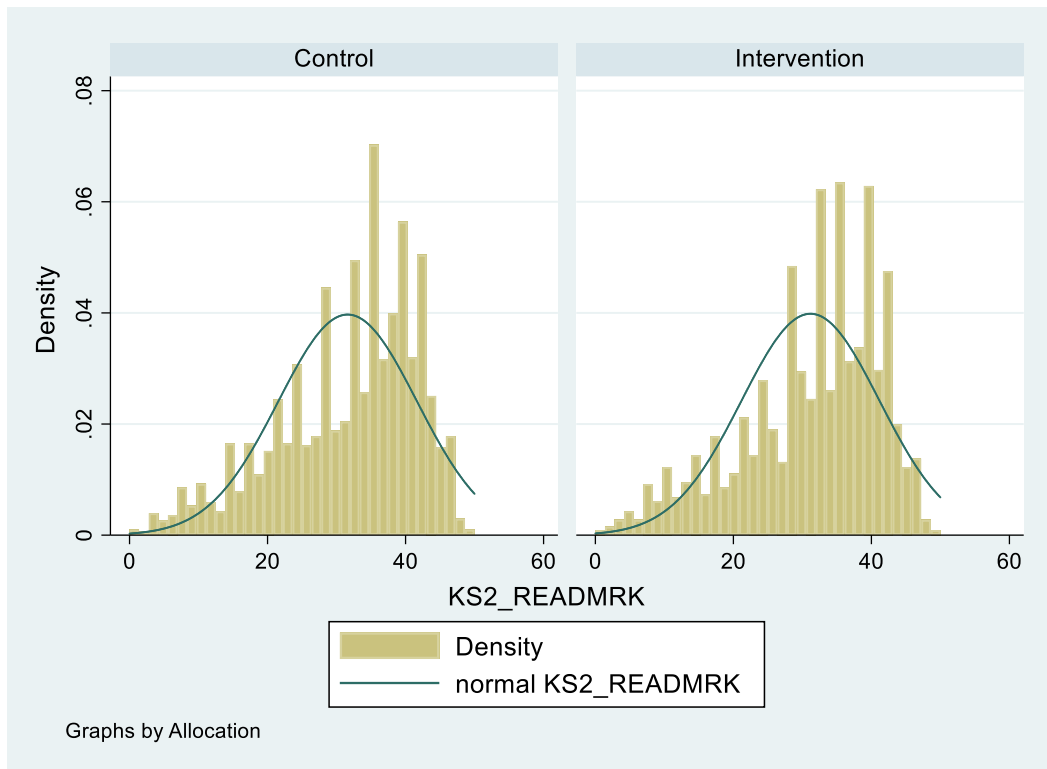
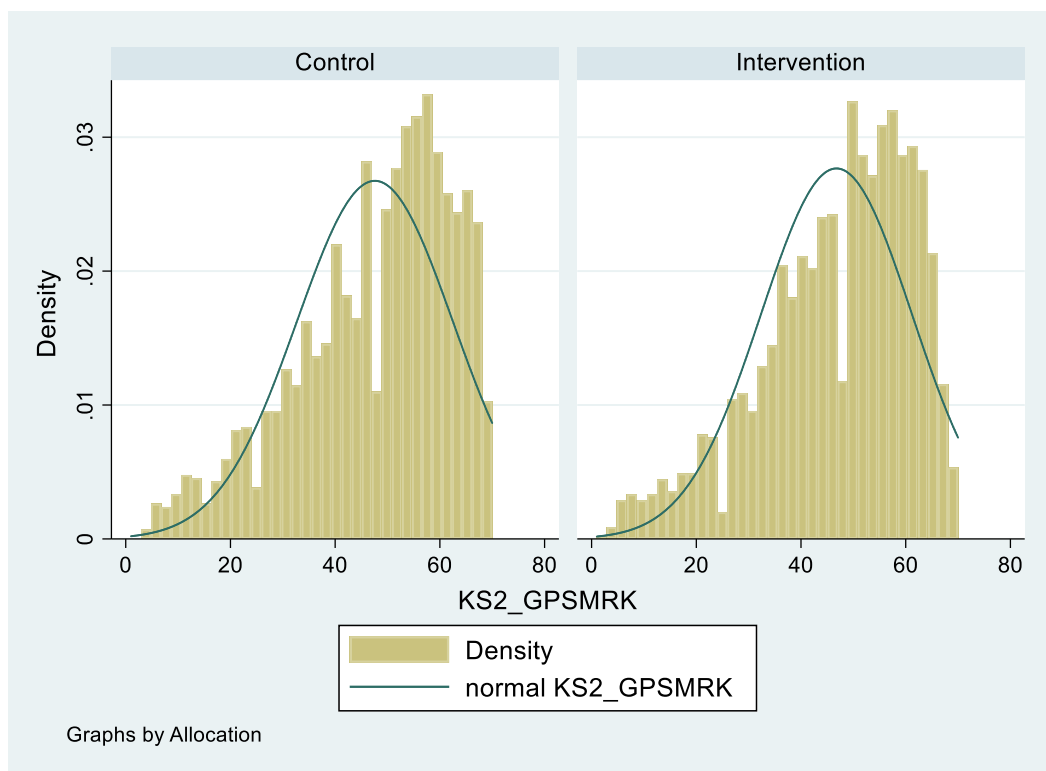


Figure D4: Figure D2: Distribution of KS2 GPS (secondary outcome)



Appendix E: Primary analyses and calculation of effect size

This Appendix provides additional model details for the intention-to-treat (ITT) impact analysis of Integrating English on KS2 writing attainment. The Appendix also explains how the model coefficient was converted into the effect size statistics shown in the main report. At the end of the Appendix is an example of the Stata code that was used for these analyses. Full model details and Stata Do files can be provided on request.

Table E1: Headline ITT analysis of KS2 writing attainment

	Null Model		Model 1.1		Model 1.2		Model 2		
	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.	p=
Group (Intervention)	-	-	-0.05	0.595	-0.19	0.540	-0.29	0.518	0.577
EAL Fluency in English Scale (Reference = not EAL)									
<i>A (New to English)</i>	-	-	-	-	-	-	-9.84	7.360	0.181
<i>B (Early Acquisition)</i>	-	-	-	-	-	-	4.18	3.818	0.273
<i>C (Developing Competence)</i>	-	-	-	-	-	-	3.17	1.800	0.078
<i>D (Competent)</i>	-	-	-	-	-	-	-1.41	1.898	0.456
<i>E (Fluent)</i>	-	-	-	-	-	-	0.31	1.676	0.855
Y5 classes per school (Reference =1)									
2	-	-	-	-	-	-	-0.06	0.763	0.937
3 or more	-	-	-	-	-	-	0.76	0.825	0.355
Hub area (Reference = South East)									
West Mids	-	-	-	-	-	-	0.84	0.816	0.306
North	-	-	-	-	-	-	0.68	0.865	0.432
Camb/Norfolk	-	-	-	-	-	-	-1.15	0.745	0.124
Leics/Pboro	-	-	-	-	-	-	0.87	0.952	0.362
Baseline/Pre-test Attainment (centred)									
Pupil level	-	-	0.32**	0.007	0.32**	0.007	0.32**	0.007	<0.001
School level	-	-	-	-	0.28**	0.067	0.36**	0.070	<0.001
<i>Constant</i>	15.71	0.299	15.56	0.416	15.61	0.377	15.33	0.804	<0.001
<i>Variance decomposition (s.e.)</i>									
School Level	6.35	1.140	6.56	1.125	5.28	0.919	4.29	4.30	0.763
Pupil Level	30.81	0.734	19.84	0.473	19.84	0.472	19.66	19.84	0.472
Total	37.17		26.40		25.12		23.95	24.14	
<i>School Level ICC</i>	0.17		0.25		0.21		0.18		
<i>Explanatory Power</i>									
School Level	-		-0.032		0.168		0.324		
Pupil Level	-		0.356		0.356		0.362		
Total	-		0.290		0.324		0.366		
<i>Hedges g effect size</i>									
95% CI (Upper)	-		+0.18		+0.14		+0.12		
95% CI (Lower)	-		-0.20		-0.20		-0.22		

*In all models, N schools = 80, pupils = 3,607. Significance ** = p<.01, * = p<.05.*

Calculating Hedges g effect size

As specified in the statistical analysis plan, the impact of Integrating English was measured using the Hedges g effect size statistic based on the formula shown below.

$$ES = \frac{(T - C)_{adjusted}}{\sqrt{\delta_s^2 + \delta_p^2}}$$

Where δ_s^2 is the school level variance and δ_p^2 is the pupil level variance for the (stage 0) empty model and $(T - C)_{adjusted}$ is the coefficient estimate for the group identifier dummy variable from the (full) model.

From Table E1, the total variance ($\delta_s^2 + \delta_p^2$) is 37.17 and so the standard deviation $\sqrt{\delta_s^2 + \delta_p^2}$ is 6.096.

Also from Table E1, $(T - C)_{adjusted}$ for the full model is shown as -0.29.

$$\mathbf{Hedges\ g} = \frac{(T - C)_{adjusted}}{\sqrt{\delta_s^2 + \delta_p^2}} = \frac{-0.29}{6.096} = -0.0476 \sim \mathbf{-0.05\ standard\ deviations}$$

The upper and lower 95% confidence intervals for the coefficient are similarly divided by 6.096 to convert them into confidence intervals for the effect size.

Example of STATA Code used in multilevel analyses for Table E1

Null Model Stage 0: mixed KS2_Writing || SchoolID:

Simple Model 1: mixed KS2_Writing Group Pupil_baseline || SchoolID:

Simple Model 2: mixed KS2_Writing Group Pupil_baseline GLPiE_School || SchoolID:

Full Model: mixed KS2_Writing GroupXX
 Fluency_A Fluency_B Fluency_C Fluency_D Fluency_E
 b1.Number_Classes b1.Hub_area Pupil_baseline GLPiE_School || SchoolID:

Appendix F: IPE survey

This follow-up survey is part of the evaluation of *Integrating English*, a programme delivered by Enfield Council to improve the English language proficiency among school pupils. It should be completed by all Y5 and 6 teachers. The survey should take around 10 minutes to complete.

Integrating English is funded by The Education Endowment Foundation, The Bell Foundation and Unbound Philanthropy. The independent evaluation is being undertaken by Sheffield Hallam University.

Data from this survey will be used to assess the effectiveness of the programme and provide insights into how EEF can best support schools to use research evidence to improve outcomes for pupils.

By completing the survey, you are giving your consent for us to use the data for research purposes. All data will be treated confidentially and will be fully anonymised in any publication arising from the research. All data will be stored securely on a password protected server at Sheffield Hallam University. No individual or school will be identifiable in any publication.

If you have any questions about this survey, or about the wider project, please direct them to the project manager, Dr Martin Culliney (integratingenglish@shu.ac.uk).

Many thanks for your help with this study.

A. Background

1 School Name:

2 Postcode:

3 What is your role in the school? Please tick all that apply:

- Headteacher
- Deputy or Assistant Head
- Literacy Coordinator
- Classroom teacher
- Other [please specify]:

4 How many years have you been teaching?

- One year or less
- 2-4 years
- 5-8 years
- 9 or more years

5 Have you undertaken any training relating to the teaching of grammar in the past 3 years?

- No (go to q6)
- Yes (go to q5a)

5a How much training did you undertake? Please tick all that apply:

- One day or less
- 2-5 days
- 6 days to 4 weeks
- More than 4 weeks

5b Please describe the training [who provided it, the focus of it] below:

B. Your views and experiences

6 To what extent do you agree with each of the following statements? Tick one box on each line:

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
Language learning plays a strong role my teaching across all subjects					
I am very knowledgeable in relation to the use of literacy across the curriculum					
I need significant additional training to improve my knowledge of teaching grammar					
I am confident that my explanations of English grammar are understood by pupils					
I usually refer to language and literacy in teaching across the curriculum					
My students engage with language issues in all subjects					
The only time I focus on grammar is in English lessons					
English language strategies that I use with students can be applied to any subject					
Language-focused planning supports my teaching across the curriculum					
I have access to high quality language-focused materials to support my teaching across the curriculum					
I regularly discuss the demands of the English language with other teachers in my school					
I regularly discuss the demands of the English language with teachers in other schools					
I know who to ask in my school if I have problems with the language demands of cross-curricular teaching					
I feel confident in my ability to deliver the English grammar goals of the language curriculum					

7 Is there a designated leader in your school to support the language demands of cross-curricular teaching?

- Yes (go to q8)
- No (go to q9)
- Don't know (go to q9)

8 How do you rate the support for the language demands of cross-curricular teaching provided by this leader? Tick one box only:

- Highly effective
- Fairly effective
- Neither effective nor ineffective
- Fairly ineffective
- Highly ineffective

C. Pupil outcomes

9 To what extent do you agree with each of the following statements? Tick one box on each line:

	All of the time	Nearly all of the time	Most of the time	About half of the time	Not very often	Rarely	Never or almost never
Overall, pupils are attentive in class							
Overall, pupils can identify the role of language in the subjects they learn about							
Overall, pupils usually require repeated explanation of classroom tasks							
Overall pupils find learning across the curriculum interesting							
Overall pupils are curious about new topics across the curriculum							
Overall, pupils find learning across the curriculum relevant to their lives							
Overall, pupils enjoy learning							
Overall, pupils behave well in class							

For question 10 (below), we would like you to consider only those EAL students now referred to as Categories B, C and D in the UK government's coding scheme.

Category B pupils are in the early stages of language acquisition. EAL Pupils are **coded as B** if they have become familiar with some subject specific vocabulary; they follow day-to-day social communication in English and participate in learning activities with support; have begun to use spoken English for social purposes; can understand simple instructions and can follow narratives or accounts with visual support; and they have developed some skills in reading and writing.

Category C pupils are developing their second language competence. EAL Pupils are **coded as C** if they require ongoing EAL support to access curriculum fully; they participate in learning activities with increasing independence; are able to express self orally in English, but structural inaccuracies are still apparent; are able to follow abstract concepts and more complex written English; they require ongoing literacy support, particularly for understanding text and writing.

Category D pupils have wide English vocabulary, but with some gaps. EAL Pupils are **coded as D** if they show solid understanding of grammar and make fewer mistakes. They read a range of texts, spell most words correctly, and write competently. Some EAL support is occasionally needed to complete complex curriculum tasks.

10 To what extent do you agree with each of the following statements? Tick one box on each line:

	All of the time	Nearly all of the time	Most of the time	About half of the time	Not very often	Rarely	Never or almost never
EAL pupils (in categories B, C and D) are attentive in class							
EAL pupils usually require repeated explanation of classroom tasks							
EAL pupils find learning across the curriculum interesting							
EAL pupils are curious about new topics across the curriculum							
EAL pupils find learning across the curriculum relevant to their lives							
EAL pupils enjoy learning							
EAL pupils behave well in class							

D. Final reflections

11 Have you changed your approach to teaching English across the curriculum in the past 12 months?
Yes [go to q11a] No [go to q12]

11a Please describe what changes you have made below

11b Please describe what motivated you to make these changes, and any support you used (such as CPD, mentoring, other resources)

12 Please describe your approach to teaching English across the curriculum below:

13 If you have any more comments about teaching English across the curriculum, please write in the box below:

Thank you for completing this questionnaire.

Appendix G: Observation schedule for training visits

Content focus of Training

Indicator	Evidence / Comments
Curriculum content that helps teachers understand how pupils learn, both generally and in specific subject areas (including subject-specific pedagogy) and enables participants to access the theory and evidence underlying the relevant pedagogy, subject knowledge, and strategies.	
There is a logical thread between the various components (training, theory, support etc.) of the programme.	
A focus on learner progression, starting points and next steps, including formative assessment, enables teachers to see the impact of their learning and work on their pupil.	
Content includes alternative pedagogies for pupils with different needs.	
Content takes account of different teachers' starting points.	

Evidence of Active Learning in the Training Sessions

Indicator	Evidence / Comments
Opportunities are provided for teachers to reveal and discuss their beliefs and experience, and test ideas from different perspectives. (This may include helping participants see how better outcomes are possible, particularly among schools where achievement has been depressed over time.)	
Teachers are engaged in analysis of and reflection around the underpinning rationale for practices changes, the supporting evidence.	

Evidence of Relevance to Teachers

Indicator	Evidence / Comments
The content and activities have overt relevance to participants' day-to-day experiences with, and aspirations for, their pupils.	

Claims Made, and Evidence Provided, for the Programme

Indicator	Claim / Evidence / Comments
The Integrating English programme effectively develops teachers' knowledge and understanding of language, based on systemic functional linguistics, in different subjects	
The knowledge and understanding from this programme leads to improved classroom practice	
The knowledge and understanding from this programme leads to improved pupils' language performance in the classroom	
The training is effective in informing how practice can be changed	
Participants are engaged during the sessions	
The key points are covered thoroughly	
The timing is suitable for the amount of content to be covered	
Overall, participants responded positively to the session	

Appendix H: Interview schedule for school visits

Questions for Teacher Interviews (post-observation)

Please read this into the tape.

We are gathering data to help us evaluate the 'Integrating English' project. This interview forms part of the evaluation. Any data collected today will be kept securely and will remain confidential. Only members of the evaluation team will have access to the anonymised transcript. Use of the interview data for this evaluation or any research project or publication will ensure anonymity. You are welcome to withdraw this consent at any time during the interview and for up to 6 days afterwards.

Do you understand what I have told you and do you consent to the use of this interview for evaluation and publishing purposes?

Please state your name:

1. Was that a fairly typical lesson with that group?
2. Were you able to achieve what you had planned? /
 - a. What language-based objectives did you plan for this lesson?
 - b. Do you think you were able to achieve these?
3. What ideas from the LiLAC training were you able to include in this lesson? What other ideas from the LiLAC training have you used in your lessons?
4. Which ideas from the training have been most useful in your teaching?
5. Have you noticed an effect on your EAL students' use of language as a result of your focus on language? Have you noticed an effect on other students' use of language as a result of your focus on language?
6. Can you identify any changes or improvements in your teaching as a result of the training?
7. What kind of support have you been able to use?
8. How useful have you found the LiLAC training and materials, the schemes of work, and the online FRONTER support since the training ended? Which have you drawn on the most to help you?
9. What kind of support do you get from your leadership team at the school to help you implement the ideas in your classroom?
10. How well have you been able to implement the scheme of work, considering the many pressures on the curriculum, the time in the classroom and other school priorities?
11. Has anything prevented you from using the ideas from the LiLAC programme?
12. Do you think your views of language and how to teach it have changed since the training?
 - a. To what extent? / Can you give me a few examples
13. Do you have any plans to share what you have learned with other teachers?
14. Do you have any further comments?

Please read this into the tape: **Thank you for your time**

Questions for School Leaders Interviews

Please read this into the tape.

We are gathering data to help us evaluate the 'Integrating English' project. This interview forms part of the evaluation. Any data collected today will be kept securely and will remain confidential. Only members of the evaluation team will have access to the anonymised transcript. Use of the interview data for this evaluation or any research project or publication will ensure anonymity. You are welcome to withdraw this consent at any time during the interview and for up to 6 days afterwards.

Do you understand what I have told you and do you consent to the use of this interview for evaluation and publishing purposes?

Please state your name:

1. What ideas from the LiLAC programme have you heard about from your teachers?
 - a. Which ideas, if any, are described most positively?
 - b. Which ideas seem to be most suited to EAL students?
 - c. Which parts of the programme appear to benefit all students?
2. Have you noticed any discussion between teachers in this school or with teachers in other schools about this project?
3. What steps have been taken to help teachers discuss with each other?
 - a. Are there any plans to take further steps to enable networking between teachers on this project?
4. How would you describe the level of enthusiasm amongst the teachers on this project?
5. What actual or potential conflicts have you noticed between the project aims and the smooth running of the school?
6. Have you noticed any changes in the LiLAC project's teachers' classes on:
 - a. pupil engagement in lessons?
 - b. pupil behaviour in lessons?
 - c. pupil confidence with language?
7. How much do you think the supported schemes of work have contributed to the implementation of the programme?
 - a. Could you suggest further types of support?
8. Do you have any further comments?

Please read this into the tape: **Thank you for your time**

Time and Date		School			
Subject / curriculum area		Topic / theme		L.O.s	
Description of classroom				Layout	
People in room					
Lesson design / structure					
Area of investigation		Evidence			
Language-based task in class					
Description or explanation of language in class					
Task or description or explanation of particular value to EAL students					
Task or description or explanation of value to ALL students					
Comments					

Time and Date		School			
Area of investigation		Evidence			
Teacher offers clear description or explanation of function of language					
Teacher uses 'genre pedagogy' technique or stage in class					
Teacher's confidence in class not affected by need to discuss language					
Language-based tasks likely to lead to better pupil understanding or use of language					
Students engage in activity					
Students remain on task and activity does not result in lack of discipline					
Students start work on tasks efficiently and complete tasks confidently					
Comments					

Appendix I: Memorandum of Understanding

Integrating English: Memorandum of Understanding for schools

The Integrating English project seeks to determine whether the LiLAC approach to English language teaching and learning can improve academic performance at Key Stage 2. The impact evaluation is based on a randomised controlled trial and the process evaluation will entail informal observations of teacher training and classroom lessons to track implementation on classroom practice along with a review of relevant teaching materials. The evaluation runs from Autumn 2016 to Summer 2018, and is due to report in 2019. Please read through the following information and sign and return one copy of this document if you wish to participate.

We hope that your school will take the opportunity to be one of 100 schools in England to participate in this research led by Sheffield Hallam University. The initiative is being delivered by Enfield Local Authority with support from Education Endowment Foundation (EEF), the Bell Foundation and Philanthropy Unbound.

Participating schools will ideally be minimum two-form entry. Schools that are larger will be eligible if they are able to commit two or more of their Years 5 and 6 teachers to the necessary training. Schools will also need to meet the technical requirements set out in the memorandum of understanding.

The project is a randomised controlled trial – the 100 schools selected will be randomly split into 50 *treatment schools* and 50 *control schools* in November 2016.

The Y5 and Y6 teachers from the treatment schools will receive the accredited professional development training, LiLAC and additional support to embed the pedagogy into the schemes of work. This will take place during the Spring and Summer terms of academic year 2016/7 and the Autumn term of 2017/8. Control schools will receive a financial incentive to participate in the trial and will have the opportunity to purchase the LiLAC training to use in their schools once the evaluation is completed as well as access to some adapted schemes of work. All schools will receive feedback on the outcomes of the study to inform future practice and will be classed as EEF partner schools.

To take part, schools will need to provide Sheffield Hallam University with data on teachers who will be taking part as well as the Unique Pupil Numbers (UPNs) and other pupil data fields for all Y5 pupils in the school. Schools will also need to inform parents and give them the choice for their child to opt-out of the baseline testing using the letter provided.

Sheffield Hallam University will administer a baseline test in November 2016. This will be the GL assessment Progress Test in English. Sheffield Hallam University will send packs of test papers for schools to distribute to pupils in class, the completed test papers will then be collected back by Sheffield Hallam and sent on to GL assessment for marking. At the end of the trial, Sheffield Hallam will retrieve KS2 data from the National Pupil Database, as this is the main outcome measure to be used in this evaluation. Sheffield Hallam will use scores from the baseline tests and KS2 examinations to gauge the effectiveness of the programme.

MOU Timetable

	Treatment schools	Control schools
Deadline for participation: 10 Oct 2016	Sign and return MOU (this document). Provide pupil data to Enfield, who will then pass on to SHU	Sign and return MOU (this document). Provide pupil data to Enfield, who will then pass on to SHU
31 Oct-11 Nov 2016	Pupil baseline testing Pre-intervention teacher survey	Pupil baseline testing Pre-intervention teacher survey
Spring term 2016/17	LiLAC training for Y5 teachers, adapting schemes of work and informal observations of teacher CPD training	Business as usual
Summer term 2016/17	Y5 pupils taught using LiLAC pedagogy, Y6 teachers receive training. SHU to assess schemes of work in all schools and review data collected by development team	Business as usual
Early Autumn term 2017	Y6 teachers complete outstanding training where necessary (must be finished by Autumn half term)	Business as usual
From Autumn half term to Summer 2017/8	Y6 pupils taught using LiLAC pedagogy	Business as usual
Spring term 2017/8	Post-intervention survey for teachers. Case studies in 15 schools comprising school visits, evaluation of schemes of work and classroom materials; informal observations of delivery; and interviews with two teachers and one mentor in each school as a minimum. Additional telephone survey of leads in non-case study schools. Review and analysis of evaluation data generated by development team	Post-intervention survey for teachers Schools will be paid £200 for taking part
Summer 2018	Key Stage 2 testing as normal	Key Stage 2 testing as normal

Prior to the trial starting ALL schools will agree to:

- Provide information on request about the school and the teachers who will be involved in the project. This includes confirming teachers' attendance at CPD and all current activities related to provision for EAL pupils.
- Issue information about the project and opt-out consent forms to parents and provide details of any parents opting for their child not to participate.
- Allow the evaluation team to issue pre- and post-intervention survey to teachers involved in the project.
- Provide Unique Pupil Numbers (UPNs) plus other data for the pupils who are involved in the trial and agree that the research team can access the National Pupil Database for KS2 data as well as demographic data such as free school meals status and gender.
- Provide data directly to the evaluation team on pupil EAL status, indicating whether each pupil is classified as EAL, and their score on the five-point EAL scale becoming mandatory from September 2016.
- Allow the GL PTE test to be administered during November 2016.
- Record any issues that might affect the fidelity of the implementation of the intervention, such as any changes in the teachers who deliver the interventions.

Please note that no individual school or pupil will be named in any report or publication arising from the research.

In addition to the above, schools selected as TREATMENT SCHOOLS will agree to:

- Appoint a project lead in the school who will function as the point of contact for the project, will support and ensure teacher engagement, facilitate school based sharing between Y5 and Y6 teachers and attend 3 meetings.
- Support participating teachers to complete the accredited LiLAC training (4 days plus networking support days), during 2017/18 as per the timetable above.
- Develop and share selected cross-curricular schemes of work with support and use these as the basis of teaching in Y5 and Y6.
- Allow members of the evaluation team to visit and informally observe lessons at pre-arranged convenient times.
- Allow participating staff to take part in, for example, research interviews, surveys and events as required by the project within reasonable scope of their time and availability.
- Not share project materials with other schools during the evaluation period.

The PROJECT will make the following commitments to schools selected as treatment schools:

CPD and student/teacher materials

- Provide the LiLAC CPD training to all participating teachers.
- Provide LiLAC manuals and materials as appropriate for the LiLAC training. Provide support opportunities for adapting schemes of work.
- Provide additional support, networking and training sessions for all teachers.
- Provide access to the online forum Fronter for support and sharing of schemes of work.

Ethics/ Data protection

- Perform all necessary checks to ensure that all school staff and project researchers act in accordance with ethical procedure. All researchers entering school premises will hold current DBS (formerly CRB) clearance.

- Test results and pupil data will be treated with the strictest confidence. The tests results will be matched with data from the National Pupil Database (including if available item-by-item scores) and potentially other government data sets, and shared with researchers at Sheffield Hallam University, the Education Endowment Foundation's data archive and the UK data archive for research purposes. We will not use pupil name or the name of the school in any report arising from the research.
- We will provide outcomes of the baseline tests so that teachers can use the results to decide how they can best support the learning of pupils.
- LiLAC project resources are owned by the government of South Australia. Enfield Local Authority is responsible for the delivery of this training as part of this programme. Unless stated otherwise, South Australia is the owner of all intellectual property rights in the materials. These works are protected by copyright laws and treaties around the world. All such rights are reserved.

If the above terms are acceptable, please sign (electronic signature or scanned signature) and date both copies, keeping one copy for your records and returning the other to integratingenglish@enfield.gov.uk by 10th October 2016.

Name	Mike Coldwell	Date	
Role	Sheffield Hallam University, Project Director	Signed	
Name	Michelle Stanley	Date	
Role	Enfield Council, IE Project Lead	Signed	
Name		Date	
Role	Headteacher (please state school)	Signed	

Appendix J: Fair Processing Notice



Integrating English (Integrating English) Evaluation

Sheffield Hallam University Fair Processing Notice

Introduction

From 25 May 2018 the General Data Protection Regulation (GDPR) will replace the Data Protection Act and govern the way that organisations use personal data. Personal data is information relating to an identifiable living individual.

Transparency is a key element of the GDPR and this Data Protection Statement is designed to inform you about:

- how and why Sheffield Institute of Education (SloE) at Sheffield Hallam University (SHU) will use your personal data
- what your rights are under GDPR, and
- how to contact us so that you can exercise those rights.

Data Subject Rights

One of the aims of the General Data Protection Regulation (GDPR) is to empower individuals and give them control over their personal data. The GDPR gives you the following rights:

- the right to be **informed**
- the right of **access**
- the right to **rectification**
- the right to **erase**
- the right to **restrict** processing
- the right to **data portability**
- the right to **object**
- rights in relation to **automated decision making and profiling**

For more information about these rights please see <https://ico.org.uk/for-organisations/guide-to-data-protection/principle-6-rights/> <https://www.shu.ac.uk/about-this-website/freedom-of-information/subject-access-request>

You can contact us if you would like to:

- request copies of your personal data held by SHU (**a subject access request**)
- exercise **your other rights** (e.g. to have inaccurate data rectified, to restrict or object to processing)
- **query** how your data is used by SHU
- report a **data security breach** (e.g. if you think your personal data has been lost or disclosed inappropriately)
- **complain** about how SHU has used your personal data

Details of who to contact are provided at the end of this document.

Why are we processing your personal data?

It is necessary for SHU to process your personal data in order to evaluate the impact of the Integrating English intervention, which is being funded by the Education Endowment Foundation (EEF). This will help to strengthen the evidence base about language teaching in schools in order to inform future policy development.

After the evaluation with EEF is complete, SHU will retain your data for research and knowledge-exchange purposes, including presentations at professional or academic conferences, or publications in professional or academic journals, for a period of ten years after the last publication arising from the evaluation. After this period, SHU will review the longer-term archival value of the data.

What is the legal basis for processing activities?

SHU are Data Controllers for the Integrating English evaluation. The processing of personal data through the Integrating English evaluation is defined under GDPR as a specific task in the public interest. The legal basis for processing your personal data is 'Public Task' (Article 6 (1) (e)).

In the production of professional or academic publications or presentations, all data will be fully anonymised and no individual or school will be identified or identifiable. Should we wish to present or publish any information where you or your school may be identifiable, for example an exemplar case study of how a school has improved as a result of participation in the Integrating English project, we will seek your consent to process your personal data.

Which Personal Data will we collect and use?

In order to provide our services we need to collect and use your personal data. Below is a list of what this includes:

Type of personal data	Pupil	Teacher
Names	X	X
Personal characteristics data including: Pupil Name, UPN, FSM status	X	
Contact details		X
Attitudinal survey responses		X
Attitudinal interview responses		X
Progress, achievement and attainment data held by your school, and in the National Pupil Database	X	
Baseline and outcome test scores	X	
Special category personal data (EAL status)	X	

We have gathered attitudinal data directly from you. Using the information we receive from you or the project provider we will also source data from national databases such as the Schools Workforce Census and the National Pupil Database.

Who will we share your data with?

The privacy of your personal data is paramount and will not be disclosed unless there is a justified purpose for doing so. Your data may be shared between SHU and:

- **EEF** for the purposes of research and evaluation.
- **GL Assessment** marked the baseline assessment papers and will have access to pupil names and ID numbers as part of this process.
- **QA Associates** will mark the outcome test papers and will have access to pupil names and ID numbers as part of this process. All personal data will be destroyed once this task is completed.
- **Randstad** will provide invigilators for the outcome test. SHU will share the names of participating schools, teachers and contact details for the purpose of arranging invigilation visits.
- **Transcribers** to produce transcripts of audio recordings of interviews and focus groups. SHU will ensure that appropriate contracts and/or data-sharing agreements are in place and that the transcribers process personal data in accordance with the GDPR and other applicable legislation.

SHU will never sell personal data to third parties

Security

SHU takes a robust approach to protecting the information it holds. This includes the installation and use of technical measures including firewalls and intrusion detection and prevention tools on the SHU network and segregation of different types of device; the use of tools on University computers to detect and remove malicious software and regular assessment of the technical security of SHU systems. SHU staff monitor systems and respond to suspicious activity. SHU has Cyber Essentials certification.

Alongside these technical measures there are comprehensive and effective policies and processes in place to ensure that users and administrators of SHU information are aware of their obligations and responsibilities for the data they have access to. Access to project data is restricted to the evaluation team and administrators associated with the project. Any sharing of the data with other researchers would require approval by the SHU Faculty of Development and Society ethics committee who will ensure that all data protection requirements are met. Training is provided to new staff joining SHU and existing staff have training and expert advice available if needed.

Retention

For the purposes of research and knowledge exchange, the retention period for your personal data will be 10 years after the last publication arising from the evaluation. After this period, SHU will review the longer-term archival value of the data. For the purposes of ongoing research, SHU will use the legal basis of 'Public Task' (Article 6 (1) (e)).

Further Information and Support

For further information about how SHU use personal data see:

<https://www.shu.ac.uk/about-this-website/freedom-of-information/data-protection-policy-statement>

The Information Commissioner is the regulator for GDPR. The Information Commissioner's Office (ICO) has a website with information and guidance for members of the public:

<https://ico.org.uk/for-the-public/>

If you have a concern about the way this project processes personal data, we request that you raise your concern with the project team in the first instance.

Contact details

SHU

Dr Martin Culliney (SHU Integrating English Project Manager)

Research Fellow, SloE, Sheffield Hallam University, S1 1WB

M.Culliney@shu.ac.uk

Professor Mike Coldwell (SHU Integrating English Project Director)

Head of Research Centre, SloE, Sheffield Hallam University, S1 1WB

M.R.Coldwell@shu.ac.uk

OR

Governance Services

City Campus

Howard Street

Sheffield

S1 1WB

foi@shu.ac.uk

0114 225 5555

If you have an ongoing concern, you can contact the Information Commissioner's Office, the body responsible for enforcing data protection legislation in the UK, at <https://ico.org.uk/concerns/>

Appendix K: Fidelity measures

Allocation		Writing	KS2 Reading	KS2 GPS
No pass	Mean	33.66	52.22	78.67
	N	64	64	64
Some pass	Mean	26.58	39	63.08
	N	286	284	285
All pass	Mean	32.09	48.29	74.42
	N	1436	1437	1439
SOW grade				
0	Mean	15.02	31.21	46.44
	N	681	1066	1067
1	Mean	18.25	29.91	47.96
	N	236	279	278
2	Mean	15.64	31.63	46.19
	N	433	493	494
3	Mean	15.52	31.36	47.41
	N	467	510	509
Action plan grade				
0	Mean	15.04	31.1	46.32
	N	562	840	840
1	Mean	16.45	32.23	49.13
	N	378	441	440
2	Mean	15.43	30.46	45.53
	N	514	591	590
3	Mean	16.4	31.22	46.97
	N	363	476	478

Appendix I: Baseline survey responses

Survey question		Disagree	Neither	Agree	N
Q6_1 - Language learning plays a strong role in my teaching across all subjects ($\chi^2 = 1.081$, $df=2$, $sig. = .583$)	INT	5.2%	10.3%	84.5%	58
	CONT	1.7%	10.2%	88.1%	59
Q6_2 - I am very knowledgeable in relation to the use of literacy across the curriculum ($\chi^2 = 3.386$, $df=2$, $sig. = .184$)	INT	3.4%	19.0%	77.6%	58
	CONT	0.0%	11.9%	88.1%	59
Q6_4 - I am confident that my explanations of English grammar are understood by pupils ($\chi^2 = 2.026$, $df=2$, $sig. = .363$)	INT	3.4%	17.2%	79.3%	58
	CONT	3.4%	8.5%	88.1%	59
Q6_5 - I usually refer to language and literacy in teaching across the curriculum ($\chi^2 = 2.001$, $df=2$, $sig. = .368$)	INT	0.0%	6.9%	93.1%	58
	CONT	3.4%	6.8%	89.8%	59
Q6_7 - The only time I focus on grammar is in English lessons ($\chi^2 = 3.901$, $df=2$, $sig. = .142$)	INT	73.7%	12.3%	14.0%	57
	CONT	86.4%	3.4%	10.2%	59
Q6_8 - English language strategies that I use with students can be applied to any subject ($\chi^2 = 1.217$, $df=2$, $sig. = .544$)	INT	5.2%	15.5%	79.3%	58
	CONT	1.7%	13.6%	84.7%	59
Q6_13 - I feel confident in my ability to deliver the English grammar goals of the language curriculum ($\chi^2 = 3.797$, $df=2$, $sig. = .150$)	INT	8.6%	17.2%	74.1%	58
	CONT	3.4%	8.5%	88.1%	59

Appendix J: Pupil writing assessment (primary outcome)

Year 6 Pupil Writing 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** for planning. You may start your writing as soon as you have finished planning.

Guest Appearance

Your headteacher wants to invite someone well-known into the school to work with pupils for a day.

The person could be a children's writer, a sports person, a chef or someone else suggested by pupils.



Your task is to write a letter to your headteacher to explain who you would like to invite and persuade your headteacher why this is a good idea.

Your headteacher wants to know the name of the person you would like to come and why she or he would be a good choice.

You could choose a real person or an imaginary character.

Planning

Think about:

- how to start your letter
- what interesting details to include
- how to end your letter

Your choice of guest (**please choose one option**):

Children's writer

Sportsperson

Chef

Your own choice _____

What the person could do on their visit:

Words and phrases to persuade your head teacher about your choice:

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

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

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

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

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



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

<https://educationendowmentfoundation.org.uk>

 [@EducEndowFoundn](https://twitter.com/EducEndowFoundn)

 Facebook.com/EducEndowFoundn