YOUNG JOURNALIST ACADEMY

Evaluation Report

September 2021

Jake Anders, Nikki Shure, Dominic Wyse, Matthew Barnard, Johanna Frerichs, Amber Evans, Kimberly Bohling

THE BEHAVIOURAL INSIGHTS TEAM

UCL Institute of Education

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.

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 evaluators

The project was independently evaluated by a team from University College London (Dr Jake Anders, Dr Nikki Shure, Professor Dominic Wyse, Professor John Jerrim, Professor Gemma Moss, Professor Andrew Burn) and the Behavioural Insights Team (Dr Matthew Barnard, Kimberly Bohling, Johanna Frerichs, Amber Evans).

The lead evaluator was Dr Jake Anders.

Contact details:

UCL Centre for Education Policy and Equalising Opportunities
Institute of Education
20 Bedford Way
London WC1H 0AL

Email: cepeo@ucl.ac.uk
Telephone: +44 (0)20 7612 6000

Executive summary

The project

Young Journalist Academy (YJA) sets up 'newsrooms' in primary schools to create interest in journalism and improve pupils' writing skills. The programme typically targets pupils in Year 5. YJA is delivered by Paradigm Arts, and many of the team are current or former journalists.

YJA staff (or mentors, as they are called) come to schools for two days to 'build the newsroom'. During this time the pupils are trained and then develop and lead their own newsrooms in their schools. The YJA mentors visit schools for six more days over the school year to support pupils to produce journalistic outputs in various forms. Each visit lasts for a full day. Class teachers assist the YJA mentor during the training days, and support pupils as part of the programme. YJA mentors provide the technical equipment required for these sessions, which includes cameras, audio recording equipment, laptops for editing, and microphones. Content that pupils produce during the school year is sent to the YJA team and they publish it on their website, which receives 20,000 visitors each month.

The trial was designed as a two-arm school-level, cluster randomised efficacy trial. It ran from November 2018 to June 2019. In total, 2,137 pupils from 82 schools took part. The trial evaluated the impact of YJA on writing attainment, writing self-efficacy, and writing creativity (ideation), although we note the challenges inherent in capturing pupils' writing attainment at scale. An implementation and process evaluation (IPE) was conducted alongside the trial. This included interviews with senior leadership teams (SLTs), teachers and mentors, as well as observations with students.

This evaluation was jointly funded by the EEF and the Royal Society of Arts.

Key Conclusions
1. Children in schools that participated in YJA made the equivalent of two months' less progress in writing, as measured by the Writing Assessment Measure (WAM), on average, compared to children in other schools. This is our best estimate of impact which has a moderate to high security rating. However, as with any study, there is uncertainty around the result: the possible impact of this programme ranges from four months' less progress to positive effects of one additional month of progress.
2. Among pupils eligible for free school meals (FSM), those in schools that participated in YJA made the equivalent of three months' less progress in writing, on average, compared to children in other schools. These results have lower security than the overall findings because of the smaller number of pupils in this group.
3. There is no evidence that YJA had an impact on writing self-efficacy or writing creativity (ideation) as measured by the Writing Self-Efficacy Measure (WSEM).
4. Findings from the IPE indicated that teachers perceived the programme to have a positive impact on pupils. Of the teachers surveyed, 69% thought that YJA had a positive impact on pupils' writing. However, some teachers were uncertain about whether YJA improved writing attainment, though these teachers said the programme may have had more of an impact on engaging more reluctant writers and increasing writing confidence.
5. Among teacher survey respondents, 74% thought that YJA had a positive impact on pupils' engagement with culture and the wider world, and that there was some evidence that longer lasting changes were taking place for pupils in relation to media engagement and skill development. Some teachers also felt that YJA had a positive impact on pupils' confidence, and that the programme improved communication skills. That said, some teachers found it difficult to reconcile the amount of time required for YJA with teaching the school curriculum, and found it challenging to further embed the programme outside of the sessions.

EEF security rating

These findings have a moderate to high security rating. This was an efficacy trial, which tested whether the intervention worked under developer-led conditions in a number of schools. The trial was a well-designed two-armed randomised controlled trial. The trial was well powered but 24.5% of the pupils who started the trial were not included in the final analysis, either because schools declined to participate in testing at the end of the intervention or because pupils were absent at the point of testing.

Additional findings

Pupils in YJA schools made, on average, two fewer months' progress than those in the control group equivalent. This is our best estimate of impact, which has a moderate to high security rating. As with any study, there is always some uncertainty around the result: the possible impact of this programme also includes negative effects of four months' less progress and small positive effects of up to one month of additional progress.

The logic model for this intervention identified three mediating mechanisms:

- pupil enthusiasm and engagement with the programme;
- pupil production of media outputs;
- the teacher embedding the programme in their class outside the eight YJA programme days.

The evaluation found some evidence supporting the first two mechanisms, but less for the third.

Regarding pupil enthusiasm and engagement, the majority of surveyed teachers felt that their pupils found YJA engaging. This was due to the work of the external provider in schools and the range of different activities involved in the programme. There was also evidence of pupil production of media outputs, as the evaluation found clear evidence of children working to produce a range of media outputs. Intervention schools published, on average, 23 media items on the YJA website (range 7–46), demonstrating that children not only produced media outputs, but outputs of a publishable standard. However, regarding the third mechanism, some teachers were not actively involved in the delivery of the intervention and some teachers found it difficult to integrate the intervention into their regular teaching.

The evaluation proposes a revised logic model that includes three intermediate outcomes:

- improved ability for pupils to write with purpose;
- improved skills for writing news articles and non-fiction reports;
- improved awareness of the news and wider world.

The evaluation found evidence for each of these outcomes.

For the first intermediate outcome (writing with purpose), case study data indicated that YJA had helped pupils, particularly those described as reluctant writers, to find purpose in their written work. For the second outcome (improved writing skills), case study data showed that teachers felt pupils had developed broader writing skills and 69% of surveyed teachers thought that YJA had a positive impact on pupils' writing. Some teachers were uncertain about whether YJA would have improved writing attainment, though these teachers said the programme may have had more of an impact on engaging more reluctant writers and increasing writing confidence. For the third outcome (improved awareness of news), observation data and case studies found some evidence of changes in relation to media engagement and skill development.

Ensuring teachers have sufficient time to incorporate YJA into existing literacy lessons and activities may have improved the impact on pupil outcomes. The teacher interviews and surveys found that SLT support was important in enabling journalism-related activities outside the eight allocated teaching days. Around 20% of teachers also identified lack of technology in school as a barrier to the successful implementation of the programme.

Cost

The average cost of YJA for one class was around £943, or £13 per pupil per year, when averaged over three years.

Impact

Table 1: 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
Writing	−0.13 (−0.32, 0.05)	−2	🔒🔒🔒🔓🔓	1,613	0.16	£ £ £ £ £
Writing, FSM-eligible pupils	−0.25 (−0.51, 0.01)	−3	N/A	413	0.06	£ £ £ £ £

Introduction

Background

This evaluation is part of a round of funding between the Education Endowment Foundation (EEF) and the Royal Society of Arts, to test the impact of different arts-based learning strategies in English schools, entitled 'Learning about Culture'.[1] The aim is to improve the evidence base around arts-based education programmes. It consists of five programmes: two in Key Stage 1 (Reception and Year 1) and three in Key Stage 2 (Year 5). Despite the unique aspects of these intervention models, there are many similarities in how they are delivered and what they hope to achieve.[2] The programmes have been supported by Arts Council England.

The background for the study is that a focus on increasing attainment in literacy and numeracy has been criticised for leading to a marginalisation of art, music and cultural studies in English schools (Neelands et al., 2015). The UK Government's Culture and Sport Evidence review (Newman et al., 2010), which summarised much of the observational and qualitative research in this area, showed student participation in cultural learning programmes (from piano training to theatre-based drama projects) to be correlated with higher levels of achievement in mathematics and literacy / English in both primary and secondary school. The review also linked participation in cultural learning programmes to faster language development in the early years, and improved cognitive ability. Additionally, large cohort observational studies in the US have suggested that the mathematics and literacy gains associated with cultural participation are particularly large for students from low-income groups (Catterall, 2009, 2012).

YJA is an intervention that establishes journalism programmes or 'newsrooms' in primary schools. It has been created by Paradigm Arts[3] to foster interest in journalism, which may be a key way to improve writing skills. A meta-analysis of interventions to improve writing in pupils (Years 4–12) found positive support for the following types of interventions: strategy instruction, summarisation, peer assistance, setting product goals, word processing, sentence combining, inquiry, prewriting activities, process writing approach and study of models (Graham & Perin, 2007). YJA operates within many of these domains, but has not yet been formally evaluated.

Building on this limited, prior evidence base, this evaluation was designed to estimate the effect of participating in the YJA over the course of one school year on pupils' writing skills. This trial was designed as a two-armed clustered randomised trial, with randomisation occurring at the level of the school. This level of randomisation was selected since entire classes participate in the programme and thus the risk of contamination within schools is very high. The two arms were: (1) Participation in YJA (Treatment); and (2) Business as usual (Control). Blocking was used in the randomisation to improve cross-arm comparability of schools, to improve precision of estimates, and to allow schools that sign up early to receive their allocation sooner than they otherwise would (this is important because of the nature of the intervention, requiring that as much notice of allocation as possible be given to teachers, as it requires activity outside of their normal working hours). More detail on the intervention is provided in the next section of this report. Initially, we had planned to look at the long-term effects of participating in the YJA after one further year, looking at results from the end of Key Stage 2 national curriculum tests in English grammar, punctuation and spelling.

We note upfront that it has been necessary for the analysis of this trial to deviate substantially from our initial plans set out in the project protocol and statistical analysis plan (SAP). These stem from issues in accessing the pre-test data that we expected to be able to obtain from the Department for Education (DfE)'s National Pupil Database (NPD). During the implementation of the trial, the DfE changed the way in which data from the NPD are made available to researchers, switching from providing extracts that can be used alongside project data within evaluators' own secure computing systems to requiring access within the Office for National Statistics (ONS) Secure Research Service (SRS). In turn, this means that it is now necessary for project data to be uploaded to the SRS. Given that this project data is considered personal data over which we, as evaluators, are data controllers, this requires the conclusion of an appropriate data sharing or processing agreement between the evaluator and the DfE and/or the ONS in order to provide legally required reassurance by the DfE/ONS about the treatment of personal data over which the evaluator is controller.

---

[1] See the RSA website for further details (https://www.thersa.org/globalassets/pdfs/reports/rsa-learning-about-culture-report.pdf)
[2] For an overarching flow diagram of the programme similarities, please see Appendix 1 in the evaluation protocol
(https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/YJA_Evaluation_Protocol_(amended).pdf).
[3] Paradigm Arts is an organisation focused on arts education. See https://www.paradigmarts.co.uk/about for more details.

Due to extended negotiations and delays between the evaluators and DfE/ONS, which we understand to have been severely exacerbated by additional workload due to the COVID-19 pandemic, in the interests of completing these evaluations and after discussion with the EEF and project teams, the decision was made to proceed with analysis with deviations from protocol flagged as we moved through the methods section. These deviations were reported to EEF ahead of conducting the analysis. Beyond issues inherent in deviating from pre-registered protocol, the main implications for the analysis are reductions in the statistical power relative to expectations.

It is important to understand the implications of this change. The purpose of including baseline measures in the current evaluation is to increase its statistical precision (i.e. to reduce the uncertainty around intervention impact estimates, which makes them more likely to be statistically significant). Importantly, both the original and the substituted baseline measures are taken from prior to the randomisation and intervention. Therefore, due to the randomised nature of the evaluation, their inclusion does not bias any intervention impact estimates, but only affects the statistical uncertainty around these estimates (i.e. the extent to which they are detectable as statistically significant). Given EEF policy to report impact estimates, whether or not they are statistically significant, there is an increased risk that headline positive or negative effects are just due to this uncertainty, rather than representing a true effect. As a result, we particularly stress the importance of statistical significance as a check on interpretation of the results in this report.

Intervention

1.   **Brief name:** Young Journalist Academy (YJA)

2.   **Why (rationale/theory):** YJA operates through setting up journalism programmes, or 'newsrooms', in primary schools. Primary school pupils receive training from YJA staff and then develop and lead their own newsrooms in their schools. They produce journalistic outputs in various forms over the course of a school year, which are published for the school and on the YJA website for a wider audience. These outputs could include print, audio or video content. The programme has been developed to stimulate interest in journalism as well as improve pupils 'writing skills and motivation for learning (see Figure 1 for a logic model of the intervention).

3.   **What (materials):** Resources and materials[4] for the intervention included (see Figure 1, *Activities*):

(a)   All materials/worksheets were provided via email and physical copies used in sessions.

(b)   Web resources were available to the teachers and pupils, including all media content produced by pupils.

(c)   Within the classroom sessions, the YJA team provided all technical equipment, which included cameras, audio recording equipment, laptop for editing (audio & video) and microphones.

4.   **What (procedures):**[5] The programme begins with a set-up meeting organised by YJA staff for the participating school. At this meeting, which includes the entire school staff, the school's SLT identifies a teacher who will take the lead on guiding the YJA over its four-phase implementation stage.

In the first phase, the YJA staff (known as the YJA mentors) come to the school for two days of training in order to 'build the newsroom'. Typically, there is one dedicated YJA mentor per school. This occurs within the classroom of the lead teacher, with the entire class taking part. During this time, pupils in the class are appointed by the YJA mentors as editorial staff, following an application process. Establishing the school-based newsroom is intended to allow for quality assurance before work is submitted to the YJA team, since the editors are responsible for reviewing the content submitted by their peers.

---

[4] See https://youngjournalistacademy.com/ for further details on resources relating to the Young Journalist Academy.
[5] The procedures described here were agreed upon at the set-up meeting and detailed in the Evaluation Protocol (https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/YJA_Evaluation_Protocol_(amended_2).pdf) and differ slightly from the usual YJA programme (e.g. the evaluation required full day sessions), details of which may be found at https://youngjournalistacademy.com/.

During phases 2 to 4 there are six more classroom days focusing on specific skills, with two days provided for each of the following activities: article writing, radio production and film/TV production. These days are delivered by the YJA mentor for the participating pupils.

Over the course of the programme, pupils are supported by their YJA mentor to take greater responsibility to run the newsroom more independently and produce journalistic content ensuring that content production continues even when YJA staff are not present. A consultation and review process is conducted by the YJA mentors with the schools at the halfway point and at the end of the school year, to monitor progress, participation and engagement. The review process is aimed at improving outcomes and providing further tailored classroom-based solutions for school cohorts (e.g. more guidance on how to produce specific types of media products).

Content that is produced by the pupils during the course of the school year is sent to the YJA team and they publish it on their website, which receives 20,000 visitors per month. All content undergoes a rigorous process of remote moderation and editorial support. All work is sent via the school, and feedback is provided by the YJA mentor if required before publication. Any amendments required from an editorial point of view must be actioned before publication can happen. The YJA mentors provide editorial input and support and ensure rapid responses to keep momentum. The checking and editorial process prevents en masse submissions, and promotes the role of editing and professionalism within the production process.

5. **Who (recipients):** YJA is targeted at Key Stage 2 (Year 5) pupils. All Year 5 pupils in a school participate in the intervention.

   For the purposes of the evaluation, it was agreed that one Year 5 teacher would be selected by each participating school's SLT. Their class would be the focus of the evaluation.

6. **Who (implementers):** YJA is delivered by Paradigm Arts. The YJA team is comprised of individuals with professional media training. Many of them are current or former journalists. The YJA mentors (as they are called) also receive training on working with young people.

   In addition, class teachers assist the YJA mentor during the delivery of the training days, and support pupils and provide guidance and feedback on their work as part of the programme (see *Activities* in the logic model, Figure 1).

7. **How (mode of delivery):** Sessions were delivered face-to-face by the YJA mentors to an entire Year 5 year group during English lessons.

8. **Where (setting):** Sessions were delivered in the Year 5 classroom. YJA mentors provided all technical equipment required for these sessions, including cameras, audio recording equipment, laptops for editing audio and video pieces, and microphones.

9. **When and how much (dosage):** The YJA intervention occurred over the course of an entire school year. There are eight dedicated training days (as outlined under point 4) with the YJA mentor. After these sessions, the participants work on content during English class time (as well as in their own time if they desire), which is sent to the YJA team throughout the school year.

10. **Tailoring:** The intervention is not planned to be personalised, meaning that all pupils receive the same intervention.

11. **Modifications:** Modifications to recipients compared to those reported in the protocol are discussed in point 5 on *Who (recipients)* above.

12. **How well (planned):** The compliance of this intervention was measured at the school level, which reflects the intervention delivery method. A school is considered to have complied if and only if the following two conditions are met:

(a)   schools must allow for all eight days to be conducted in schools with the YJA delivery team;

(b)   all schools must have uploaded at least ten media items by the end of the intervention OR be considered by mentors still to be adequately participating by fulfilling a set of criteria defined by the delivery team.

In comparison, implementation fidelity was conceptualised as how the way in which the intervention was implemented in practice compares to the intended implementation of the intervention as described in the Intervention section of this report. Case study data was used to examine the variation in how the intervention was implemented and any adaptations made in the case study schools, alongside identifying barriers and facilitators to implementing the intervention with fidelity.

To maintain or improve compliance and fidelity, the core editorial team of pupils is selected at each school (see Figure 1, *Moderating factors* and *Mediating mechanisms*). They assist the YJA team in moderating the quality of the outputs produced and maintaining the momentum of the intervention.

Figure 1: Logic model (original version)
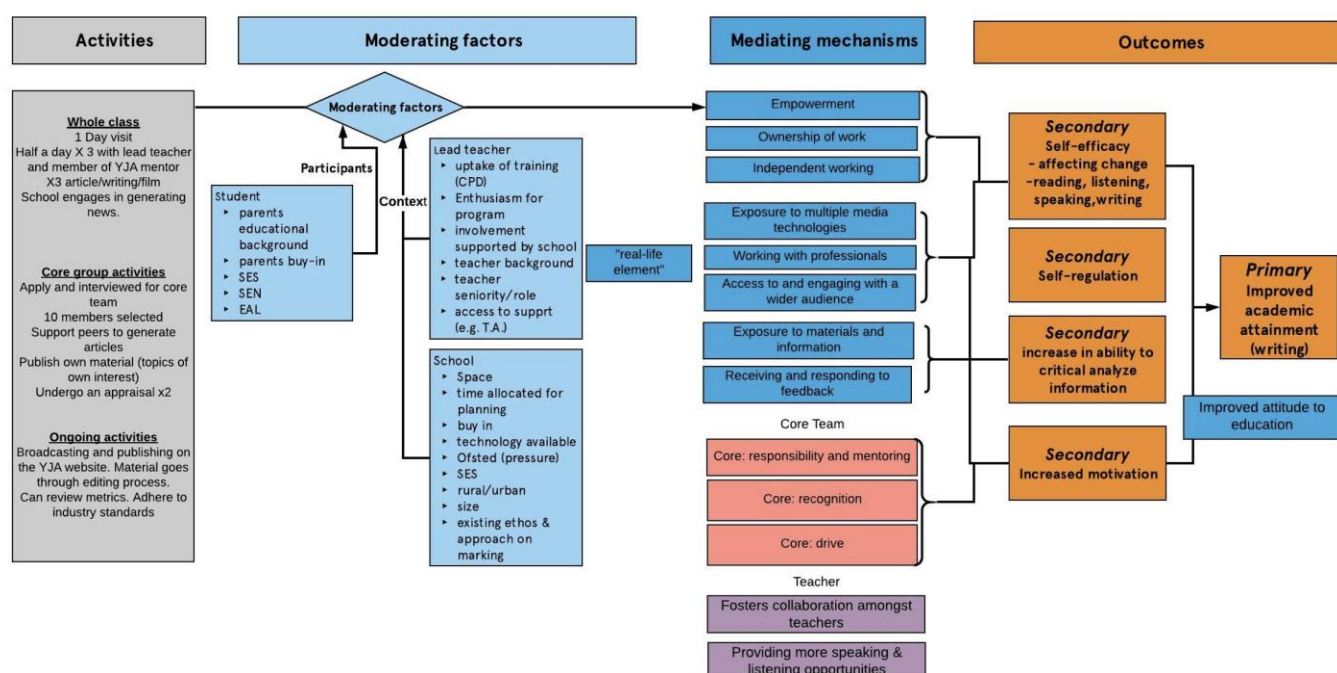
Evaluation objectives

The primary objective of this evaluation was to estimate the effect of participating in the YJA over the course of one school year on pupils' writing skills. The impact evaluation sought to answer the following research questions:

1. What is the effect of participating in the YJA over the course of one school year on pupils' writing skills? [primary research question]

2. What is the effect of participating in the YJA over the course of one school year on pupils' writing self-efficacy? [secondary research question]

3. Does participating in the YJA over the course of one school year have an impact on pupils' perception of their own capacity to generate ideas? [secondary research question]

The implementation and process evaluation (IPE) focused on addressing research questions related to the implementation, delivery and perceived impact of the intervention. The IPE was designed to explore overarching implementation questions across all five Learning about Culture trials, as well as research questions specific to YJA. The four overarching questions were written based on cross-project similarities to facilitate comparisons; however, not all questions apply to each programme due to variations in training and delivery (see Appendix R for model depicting overarching similari-

ties). The relevant overarching implementation questions that are explored across all projects, as well as research questions specific to YJA, are detailed below. Where there were deviations from the protocol, they are noted in relation to the question and described further below.

Learning about Culture overarching IPE questions

1. In what ways was the programme implemented? What are the barriers and facilitators of delivery (Fidelity)? In particular:

(a) SLT buy-in;

(b) Delivery of the intervention**:** (i) consistent across sites; (ii) whether it appears to be effective in supporting children's attainment; and (iii) whether it appears to facilitate children's engagement;

(c) Delivery of training: (i) the extent to which is it consistent across sites; and (ii) whether it appears to be effective in ensuring that teachers understand the aims and main features of the intervention. *[omitted; not applicable to YJA]*

2. To what extent did the schools engage with the intervention in line with the intervention aims? (Responsiveness).

3. How was the quality of the intervention perceived by teachers, senior leaders and teaching assistants? (Quality)

4. To what extent is the knowledge of arts practitioners delivering the intervention integrated with the pedagogic knowledge of teachers involved? (Support)

Research question (RQ) 1 was refined prior to data collection to remove the sub-question exploring the delivery of training, as this intervention did not involve any direct training of teachers.

Young Journalist Academy (YJA) specific questions

Beyond the overarching questions listed above, the IPE also sought to answer questions specific to the YJA intervention.

5. What are the mechanisms that are taking place in the intervention and to what extent are they bringing about change? (Mechanisms)

6. The relationship between the 'core' team and the rest of the class: How does this affect engagement with the programme? (Responsiveness) [Deviation: analysis focused on quality]

7. How do schools engage more broadly in showcasing journalist outputs? (Engagement)

8. Which elements of the intervention are most widely adopted and how does this affect outcomes? (Adaption/Quality)

9. To what extent is the intervention disseminated across the school? (Reach)

10. How does the YJA intervention affect literacy in the class and school? (Mechanisms)

11. To what extent do school facilities, space and technology, affect the intervention? (Implementation environment)

RQ 6 was intended to explore the relationship between the 'core team' and the rest of the class, as well as the effect of this relationship on classroom engagement with the programme. However, in initial qualitative interviews, teachers were not able to comment on the relationship, instead focusing on the team selection process and the barriers and facilitators of selection and ways of working that affected the team's ability to function. In this way, the analysis focused on quality, and not on responsiveness as intended.

These research questions were reported in the evaluation protocol (initially published in May 2018 with revisions for clarity published in April 2019: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/YJA_Evaluation_Protocol_(amended).pdf), and further details on the quantitative approach were provided in the statistical analysis plan (published in February 2019, https://educationendowmentfoundation.org.uk/public/files/Projects/YJA_SAP.pdf).

We also planned to estimate the longer term effects of participating in the YJA after one further year, looking at results from the end of Key Stage 2 national curriculum tests in English grammar, punctuation and spelling. However, it will not be possible to carry out this analysis due to the cancellation of these assessments as a result of the COVID-19 pandemic.

Ethics and trial registration

The project's aims, methods and materials were reviewed through the processes laid out by the UCL Institute of Education research ethics committee and approved on 26 March 2018. Although the application was approved, the ethics reviewers stressed the importance of ensuring ongoing pupil assent for participation in any evaluation activities throughout the research. As such, all research assistants conducting assessments with pupils verbally described the activities to the pupils using age-appropriate language, informed them all activities were voluntary, and gave an opportunity for pupils to decline to participate.

Schools were informed about the trial through initial information from the developer and formally committed to participation by signing a memorandum of understanding. A template version of this document is included as Appendix N to this report.

This trial protocol has been pre-registered at www.controlled-trials.com, and assigned an International Standard Randomised Controlled Trial Number (ISRCTN) of ISRCTN14491875.

Data protection

As part of this project, we processed pupils' and teachers' personal data. For this reason, it was important that we processed this data lawfully, following the principles laid out in the Data Protection Act 1998 (DPA) until May 2018, the EU General Data Protection Regulation (GDPR) from May 2018 until December 2020, and the UK General Data Protection Regulation from January 2021 (the project spanned these three regulatory periods). We explain the lawful basis below with respect to the GDPR, but there are equivalent regulations in the DPA for the justifications set out below.

BIT used Article 6(1)f of the GDPR as the lawful basis for processing personal data as part of this project. This is generally known as the 'legitimate interests' basis. There's a guide to that article here:

https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr*/lawful-basis-for-processing/legitimate-interests/

BIT carried out a 'legitimate interests assessment' in support of this, identifying societal benefits to this processing of personal data. Specifically, the use of pupils' and teachers' personal data as part of this research was to understand the benefits to pupils of participating in this programme in terms of their academic attainment and other related benefits. This has public benefits that BIT believes are significant in terms of understanding whether this programme has the potential to benefit children in schools across England. Without processing these data it would not have been possible to provide this quality of new evidence.

UCL used Article 6(1)e of the GDPR as the lawful basis for processing personal data as part of this project. This is generally known as the 'public task' basis. UCL has reviewed current ICO guidance available here: https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/public-task/

and has determined that this research forms part of its performance of a task in the public interest, as one of its core purposes provided for in its Charter and Statutes. This use of data was allocated the following UCL Data Protection Registration Number: Z6364106/2017/11/69 social research.

We do not believe that any of the data we processed falls within the definition of special category data under the GDPR. This would require an additional justification under Article 9(2) of the GDPR.

We informed pupils' parents of the proposed data processing and provided the opportunity to object to this. If parents objected, then the pupils' data was never passed to us by schools. If a parent chose to withdraw their child's data at a later stage, then it was destroyed. The data controllers were named in the privacy information provided as part of this project and contact details were provided in case they had any queries about the data we held about them, including provision and deletion of their data. The relevant letters and forms have been reproduced in Appendix Q.

The information provided to parents explained in clear and plain language the lawful bases for processing, the purpose to which we put the data, that they could object to this data processing and this would be respected; the contact details of the organisation and categories of data that have been processed.

Data will be kept until the end of the research project, including academic paper writing and dissemination (and certainly not longer than 10 years, in line with UCL's policy on data retention). When it is deleted, it will be securely destroyed.

Data will be shared with EEF (who funded the trial), EEF's data contractor FFT Education (who manage EEF's Data Archive) and (in a form that is truly anonymised) to the UK Data Archive. Details of this sharing were included in relevant Data Privacy Notices.

Project team

The project team comprised Rob Pitman at Paradigm Arts and Sam Atkins at c1media. The intervention was designed by the project team and was delivered by Paradigm Arts with the support of delivery partners.

The impact evaluation was led by Dr Jake Anders and Dr Nikki Shure at UCL Institute of Education with support from Prof. John Jerrim, and was led at the Behavioural Insights Team by Kimberly Bohling and Dr Matthew Barnard. Data collection was managed by Faisa Abdi, Eleanor Collerton, Camilla Devereux, Amber Evans, Louise Jones, Alex Manby, Bridie Murphy, and Juliane Wiese of BIT. Primary data collection and marking of those data were carried out by research assistants employed by BIT, drawn from finishing students at UCL Institute of Education. The implementation and process evaluation was led by Dr Matthew Barnard with support from Amber Evans and Johanna Frerichs of BIT, with input from Prof. Dominic Wyse (UCL IPE lead), Prof. Gemma Moss and Prof. Andrew Burn at UCL Institute of Education. The cost evaluation was led by Dr Matthew Barnard with support from Faisa Abdi of BIT. The evaluation design was also supported by Daniel Carr, Dr Florentyna Farghly, Dr Jessica Heal, and Dr Pantelis Solomon of BIT.

Methods

Trial design

Table 2: Trial design

Trial design, including number of arms		Cluster randomised, two arms
Unit of randomisation		School
Stratification variable(s) (if applicable)		Proportion of English as an Additional Language (EAL) students (high/low split at sample median within randomisation batch); proportion of Free School Meals (FSM) students (high/low split at sample median within randomisation batch)
Primary outcome	Variable	Writing attainments
	Measure (instrument, scale, source)	Writing Assessment Measure (WAM), 0–32 (Dunsmuir et al., 2015)
Secondary outcome(s)	Variable(s)	Writing self-efficacy Ideation
	Measure(s) (instrument, scale, source)	Writing Self-Efficacy Measure (WSEM), 16–80 (adapted from Bruning et al., 2013) Ideation captured from sub-measure of the WSEM, 5–25.
Baseline for primary outcome	Variable	Planned to be: Phonics attainment Protocol deviation: FSM status, EAL status, class FSM composition, class EAL composition
	Measure (instrument, scale, source)	Planned to be: Phonics Screening Check (DfE) Protocol deviation: 0/1 indicator of FSM eligibility status, 0/1 indicator of EAL status, proportion of class reported FSM eligible [0,1], proportion of class reported EAL [0,1]. (All derived from school reports collected ahead of randomisation; see p.18 for further details and justification).
Baseline for secondary outcome(s)	Variable	Planned to be: Personal, Social and Emotional Development skills Protocol deviation: FSM status, EAL status, class FSM composition, class EAL composition
	Measure (instrument, scale, source)	Planned to be: EYFSP Personal, Social and Emotional Development skills (DfE) Protocol deviation: 0/1 indicator of current FSM eligibility status, 0/1 indicator of EAL status, proportion of class reported FSM eligible [0,1], proportion of class reported EAL [0,1]. (All derived from school reports collected ahead of randomisation; see p.18 for further details and justification).

This trial was designed as a two-armed, stratified, clustered randomised efficacy trial with randomisation occurring at the level of the school and outcomes measured at the level of pupils. This level of randomisation was selected since entire classes participate in the programme and thus the risk of contamination within schools was very high. The two arms were as follows:

- Participation in Young Journalist Academy (Treatment)

- Business as usual (Control)

The primary outcome of interest was improvements in pupils' writing, with the secondary outcomes of their writing self-efficacy and ideation. Further information on the approach taken to measuring these is provided below.

Participant selection

The project aimed to recruit one hundred English state-funded primary schools, based on a regional criterion as well as past participation in the YJA. The geographic areas from which schools were selected included: Lincolnshire, Notting-hamshire, Derbyshire, Rutland, London and Newcastle. This region was agreed upon with the developer (Paradigm Arts) due to their location and the location of YJA mentors; it is intended to allow for recruitment across both urban and rural settings, although without claim to representativeness in this sense. In the end, 82 schools were fully recruited and included in the randomisation.

YJA is currently delivered to pupils across the primary school age range. Year 5 was chosen for the purposes of the evaluation, given the greater perceived potential to capture writing-based outcomes from pupils at older ages, but without attempting to deliver and evaluate in Year 6 classes because of the perceived tension with Key Stage 2 national curriculum tests at the end of the year. Furthermore, evaluation of Year 5 delivery was intended to allow for medium-term follow-up in those Key Stage 2 national curriculum tests, but this was ultimately impossible due to their cancellation for the relevant year group in the wake of COVID-19.

As discussed in the intervention section, it was initially planned and recorded in the evaluation protocol that all Year 5 classes in the participating school would participate in YJA, but that only one class from Year 5 would be selected for participation in the evaluation, based on school nomination (pre-randomisation). Once the trial began, this was altered in some schools for reasons of feasibility, with intervention participation generally reduced to one Year 5 class per school. Evidence from the IPE implies that some schools expressed a preference for delivering YJA in a single class-room, but others indicated they would have welcomed broader delivery. As this is based on case study schools, we are not able to generalise this point further. All children in the teacher's class participated in the trial (other than where objections to processing personal data or participating in evaluation activities were received). Except in unforeseen and unavoidable circumstances (e.g. teachers moving school), the teachers (and, therefore, pupils) who participated were selected prior to randomisation to minimise the potential for this to introduce differences between the intervention and control groups; except in this small number of cases, all data on participating pupils was collected pre-randomisation in order to assure this.

In order to be considered for participation in the evaluation, schools had to:

(i)   agree to distribute opt-out consent[6] forms to parents;

(ii)   provide student data in order to identify a consistent relevant analysis sample and to allow for linking to the Department for Education's National Pupil Database (for pre-test data and long-term follow-up);

(iii)   identify the teacher who will participate in the trial;

(iv)   cooperate with the project and evaluation teams during the trial (further details of these requirements are outlined in the Memorandum of Understanding with Schools, available in Appendix N).

The project team advertised the trial and also approached schools through their existing networks. The team aimed to recruit schools that have larger populations of individuals receiving FSMs than the national average of 15.3 percent of pupils aged 5–10 (DfE, 2016). This appears to have been successful, judging by the proportion of pupils eligible for FSM in the analysis sample.

The eligibility criteria for schools to participate were:

---

[6] Note that this is an opt-out consent from a research ethics point of view, not from a data protection point of view. We note that since the first version of this protocol was agreed the GDPR has been implemented. As such, UCL's legal basis for processing this data is now considered to be 'public task' and BIT's legal basis for processing personal data is now considered to be 'legitimate interest'. 'Consent 'is not used by either party as a basis for the processing of personal data. Nevertheless, it remains the relevant term in respect of research ethics.

- participating schools must be English state-funded primary schools (they were recruited from Lincoln-shire Nottinghamshire, Derbyshire, Rutland, London and Newcastle);

- schools had to agree to distribute study information sheets, data privacy information and data processing objection forms to parents;

- participating schools' SLT had to nominate one Year 5 teacher and their class to participate in the intervention;

- schools had to agree that, if allocated to the control group, they would continue with 'business as usual ' for the duration of the trial;

- schools had to return a signed Memorandum of Understanding, including committing to participate fully in the study – including the collection of outcome measures in summer 2019 – regardless of which trial arm they were assigned to;

- schools had to agree to allow time for each assessment phase and liaise with the evaluation team to find appropriate dates and times for assessments to take place;

- schools had to agree that teachers in both trial arms cooperate with activities for the IPE, if requested.

Outcome measures

Baseline measures

Baseline measures for this research were planned to be drawn from the DfE's NPD. All participating schools were asked to provide personal information about pupils in the participating teacher's class that would allow a reliable link to be achieved, based on current guidance from the DfE and balancing this against personal data minimisation requirements set out in data protection legislation. Using this link, it was expected that we would obtain information on pupils' performance in the Phonics Screening Check (using the NPD variable PHONICS_MARK) for the primary outcome analysis, and assessments of pupils' personal social and emotional development from the Early Years Foundation Stage Profile (aggregated scores from NPD variables FSP_PSE_G06, FSP_PSE_G07 and FSP_PSE_G08) for the secondary outcome analyses.

Due to the data access issues described above, an alternative approach was taken, with its design informed by an intention to maximise the explanatory power of our analysis model and, hence, the precision of our treatment estimates, given the data available. As such, instead of including the planned baseline measures in the model, we substitute the available demographic information that was collected about pupils ahead of randomisation (initially intended solely for the purposes of stratification/blocking as part of the randomisation process), specifically eligibility for FSM and whether the child has English as an additional language (EAL). We include these in the model as predictors themselves, and also aggregated them to the class level to produce composition variables, given evidence that school-level aggregate predictors also provide explanatory power (Bloom et al., 2007). FSM and EAL status are both known to predict academic attainment (Strand, Malmberg & Hall, 2015; Sutherland, Ilie & Vignoles, 2015) and, as such, we expected this to improve power compared to an empty model.

Nevertheless, the improvement in statistical power is still likely to have fallen short of that we would have expected from including a prior attainment measure, as was planned. It is important to understand the implications of this change. It is important to stress that there are no expected implications for bias in our impact estimates of not having our planned baseline measures – the unbiasedness of RCT estimates derives from the randomisation, not from statistically controlling for differences at baseline. Indeed, in principle, there is no need to include any baseline measures at all in the analysis to achieve an unbiased estimate from an RCT. Inclusion of inappropriate covariates in our analysis would have the potential to introduce bias – such inappropriate covariates are ones that could have been affected by the treatment, which is why we are including pupil characteristics from prior to randomisation. The main implication of this change is a reduction in statistical precision (i.e. the uncertainty around estimates that is inherent in all evaluations is likely larger in this evaluation than it would have been, which is manifested as wider confidence intervals (CIs) or, equivalently, less likely to be statistically significantly different from zero for a given size of impact estimate). Given EEF policy to report impact estimates whether or not they are statistically significant, there is an increased risk that

headline positive or negative effects are just due to this uncertainty, rather than representing a true effect that would have been the case in the presence of more explanatory power from baseline measures. As a result, we particularly stress the importance of statistical significance as a check on interpretation of the results in this report.

Primary outcome

Writing attainment

To measure the primary outcome, we used the Writing Assessment Measure (WAM) (Dunsmuir et al., 2015; Murphy et al., 2013). The WAM was developed in order to create a valid and reliable writing assessment measure, relevant within the context of the English educational system. This measure is designed to assess narrative writing in response to a written prompt, to which pupils are given 15 minutes to write. Previous evidence suggests that this measure is reliable (test–retest correlation $r = 0.82$ over 21 days with different prompts) and valid ($r = 0.786$ with Wechsler Objective Reading Dimensions – WOLD – Written Expression sub-test) (Dunsmuir et al., 2015). The WAM prompt presented to pupils is included as Appendix O. We carried out further analysis of the performance of WAM given its lack of widespread use up to this point, which is reported below.

The WAM is based on the structure and format of the WOLD Written Expression sub-test, with modified dimensions that incorporate descriptors from the national curriculum writing attainment targets, including the following seven dimensions: ideas development, organisation and planning, vocabulary, sentence structure and grammar, spelling, punctuation and handwriting. For each of these sub-scales, the pupil can receive a mark of 1–4, with 4 being the highest. The WAM is unique as an assessment because it incorporates 'ideas development', which fits well with aims of the interventions being tested as part of the Learning about Culture project. The overarching logic model for the Learning about Culture project includes increased creativity as an outcome (see YJA evaluation protocol for further details). In the case of producing journalistic content, the YJA programme also places an emphasis on creativity (i.e. experimenting with different media forms to tell a story). In support of this, we double-weighted the score on the 'ideas development' dimension. Final scores range from zero to 32 (after accounting for double-weighting).

Robust assessment of writing is challenging, particularly during primary schooling. However, the centrality of understanding improvements in writing, and hence the need for this to be the primary outcome measure, was stressed in the initial project outlines (noting that previous trials had generally focused on reading, rather than writing, adding to the rationale for funding these projects) and set-up meeting discussions with the EEF and programme teams. Use of the WAM (Dunsmuir et al., 2015) as a measure of KS2 age pupils' writing was not the initial proposal for this trial, but emerged from discussions held as part of the project set-up meetings. The WAM is an analytic measure of writing based on equal weighting of the following criteria: handwriting, spelling, punctuation, sentence structure and grammar, vocabulary, organisation and overall structure, and ideas. There are, of course, some limitations to its use, largely stemming from the fact that it is a relatively new measure and we would have, ideally, preferred to have used a measure with a longer track record. Nevertheless, we believe that it is a pragmatic measure for the context of this research. Dunsmuir et al.'s (2015) results are encouraging in terms of the measure's internal consistency (Cronbach's alpha = 0.87), inter-rater reliability (Cohen's kappa > 0.7 for all sub-scales except 'ideas', where kappa = 0.62; kappa > 0.6 is generally considered satisfactory), and test–retest reliability ($r = 0.82$). To supplement this evidence, the EEF provided funding for us to undertake a small, informal piloting of the WAM.

The Behavioural Insights Team (BIT) conducted a small-scale pilot of the WAM in October 2017, with approximately 50 pupils from one Year 5 class, one Year 6 class, and one mixed Year 5/6 class. The aim was to understand how clearly pupils understood the prompt, how much they were able to write during the time allotted, and act as a sense check of the measurement properties reported by Dunsmuir et al. (2015). Pupils were given the WAM prompt, one sheet of A4 paper, and 15 minutes to complete the task. The results of the pilot showed that pupils had little difficulty in completing this writing task but required some additional clarification on the prompt and additional paper. In addition, the measurement diagnostics remained encouraging (albeit this may have been helped by the small sample). While we were keen to maintain consistency between the WAM as implemented by Dunsmuir et al. and this work, in order to ensure that what we do know about the WAM from their work carries forward to our context, we have made some small adaptations based on concerns identified during the set-up process and from observations arising from

the small-scale pilot. These were discussed with the WAM's lead developer, whom we gratefully acknowledge as having provided helpful informal advice as part of this process. Concerns from developers were noted and discussed with mitigating actions as follows:

- Concerns that the WAM may over-weight surface features of the language. We do agree that this is a potential concern, particularly within the context of arts-based learning evaluations, while noting that we think alternatives (such as using KS2 grammar punctuation and spelling tests) would be much worse affected by this problem. In partial mitigation we have double-weighted the ideas sub-scale. This double-weighting results in an outcome distribution that is slightly less normally distributed, but not to an excessive degree (see Figure E1).

- Concerns regarding the content and face validity of the WAM given its alignment with an earlier version of the English national curriculum. The overall aims for the teaching of writing in primary schools, as specified in England's national curriculum implemented since 2014, require that teachers develop pupils' competence in 'transcription (spelling and handwriting)' and 'composition (articulating ideas and structuring them in speech and writing)' (DfE, 2013, p. 16). The WAM is an appropriate measure of writing in the context of England's current national curriculum aims.

- Concerns that aspects of the prompt may be confusing to pupils. These were identified from the pilot. Revisions were made to the introduction of the WAM to pupils, to provide increased guidance to pupils on the purpose of the writing sample we ask them to produce, given that we understand this to be normal practice for pupils of this age when taking part in a writing activity (for the WAM prompt used in this trial, see Appendix O).

- Concerns that 15 minutes is not long enough. The observations from the pilot suggested that 15 minutes is sufficient for pupils to produce a writing sample that could be meaningfully assessed. However, clearly this was only small scale. We take this concern seriously, while wishing to maintain broad alignment with previous use of the WAM. As such, in a change from the WAM as previously used, we provided five minutes of planning time at the start of the activity, during which pupils can make notes but don't begin the writing activity itself. This also helps to make the activity more familiar to pupils, in line with the previous point.

The writing tasks for this evaluation and the other two Key Stage 2 evaluations (the Craft of Writing and Power of Pictures) that were part of the EEF/RSA Learning about Culture project were invigilated and collected in summer 2019 by a team of research assistants (RAs) coordinated by BIT as a combined exercise. Since outcome data were collected as part of a single exercise and consistent (in terms of both measure and timing) across these three evaluations, we report our analysis of the data collection and measurement with pooled WAM data collected across the three projects: pooling these data allows us to increase sample size for these analyses and, hence, reduce noise and risk of small sample bias in estimates from these analyses. The same goes for consideration of the WSEM, which follows.

Data collection

RAs were kept blinded to trial arm assignment of school in order to avoid the potential for this to bias the outcome measurement; for example, by being more lenient on timing in treatment schools. A separate group of 25 RAs (17 of whom marked tasks on the YJA project), who were also blind to trial arm assignment of school (this time, in order to avoid the potential for this to introduce bias to the trial, e.g. through unconsciously being more generous in their marking of pupils in the treatment group), marked the writing exercises against the WAM scoring sheet. This blinding is important in supporting the evaluation's internal validity. The WAM scoring sheet provided detailed criteria for assigning scores on each of the seven dimensions. During training, markers were provided with examples of student writing that exemplified each rating within a given dimension and an explanation of why that sample achieved its rating. The writing samples, scores and explanations were all provided by UCL academics who have expertise using the WAM. A random sample of the tests (approximately 3%) was independently second marked by one of the other RAs during this process (a minimum of two tests per RA per day), with a correlation of 0.75 between markers in this double-marked sample. Where discrepancies arose, these were used to feed back to markers in order to improve the consistency of marking over the exercise as a whole – this continuous improvement process may inflate this correlation

over the course of the process, but it improves the reliability of the marking relative to the alternative. As a point of reference, Dunsmuir et al. (2015) report an average marker-level intra-class correlation (ICC) of 0.97 for the WAM (range 0.93–0.99 at 95% confidence interval).

Furthermore, analysis of the basic statistical properties of the overall measure is encouraging: the distribution of the scores is normal (skewness = –0.27; see also the histogram in Figure E1 of Appendix E), which suggests there were minimal issues with floor or ceiling effects, and we calculate a Cronbach's alpha across the seven marking sub-domains of 0.84, suggesting these cohere sensibly into the overall score.

Secondary outcomes

Writing self-efficacy

As highlighted in the logic model, the impact of the intervention on writing outcomes may have an effect through pupils' engagement with and motivation for writing, which may in turn have an effect on their sense of efficacy as a writer. For this reason, we consider writing self-efficacy as our secondary outcome. In addition, self-efficacy has been highlighted in EEF's review of non-cognitive skills: the evidence indicates that 'self-efficacy for a particular task is malleable and that improved self-efficacy is associated with greater persistence, interest, and performance' (Gutmann & Schoon, 2017, p.11) and that 'the best predictors of specific academic performance are self-efficacy beliefs regarding those specific academic domains' (Pajares, 1996, in Gutmann & Schoon, 2017, p.11).

To measure writing self-efficacy, we used a Writing Self-Efficacy Measure (WSEM) which was adapted from the Self-Efficacy for Writing Scale (SEWS) measure proposed by Bruning et al. (2013), in order to make it suitable for primary school pupils through some simplification (see Appendix P for the full WSEM questionnaire used in this trial). Bruning et al.'s original measure involves 16 statements capturing aspects of writing, including 'I can think of many ideas for my writing' and 'I can avoid distractions while I write', with pupils giving marks out of 100 for their self-assessment in each of these. We used slightly simplified versions of some of these statements to better suit the primary school context; in addition, we requested responses on a five-point Likert scale, rather than marks out of 100. These adaptations were based on consultation with experts in primary literacy pedagogy and were piloted through the same process as outline for the WAM above, with resulting refinements to wording of the Likert categories as the initial versions were found to encourage pupils to exaggerate their confidence as this was seen as the 'right' answer. Bruning et al. (2013) developed a multi-factor model of writing self-efficacy; however, since the intervention is not hypothesised to have a link with their specific factors (with the exception of ideation, which we discuss separately below), we used a simple aggregate of self-assessments across all 16 statements (all are positively framed so there is no need for reverse coding). As such, possible scores range from 16–80 for each child.

As with the WAM, these tasks were administered and collected in summer 2019 by a team of research assistants (RAs) coordinated by BIT, who also marked the WSEM. RAs were kept blinded to trial arm assignment of schools. Again, consistent with the WAM, we explored the statistical properties of this measure (carried out pooled with WSEM data collected concurrently for the Craft of Writing and Power of Pictures projects), given the adaptations that were made in order to use it for this project. The overall scores are rather negatively skewed (skewness = -–1.16; see also histogram in Appendix E) which could attenuate impact estimates for this outcome; we estimate a Cronbach's alpha of 0.90 across the individual items of the measure, suggesting they cohere sensibly into the overall score, and a Cronbach's alpha of 0.79 across the three sub-domains of ideation, convention and self-regulation.

Ideation

The logic model also identifies the potential for increased creativity in the pupils who have participated in their programme. To explore this, we reported differences in the 'ideation' sub-measure of the WSEM as an additional secondary outcome measure. This measure was jointly chosen with RSA and allows us to address secondary research question 2 on pupils' perception of their capacity to generate ideas. This uses the first five questions of the writing self-efficacy measure (see Appendix O for these statements) and, as such, possible scores range from 5–25 for each child. As this is a sub-domain of the WSEM as a whole, data collection details are as per that measure. Also like the WSEM, the scores are rather negatively skewed (see the histograms in Appendix E), with potential consequences for attenuating impact estimates for this outcome.
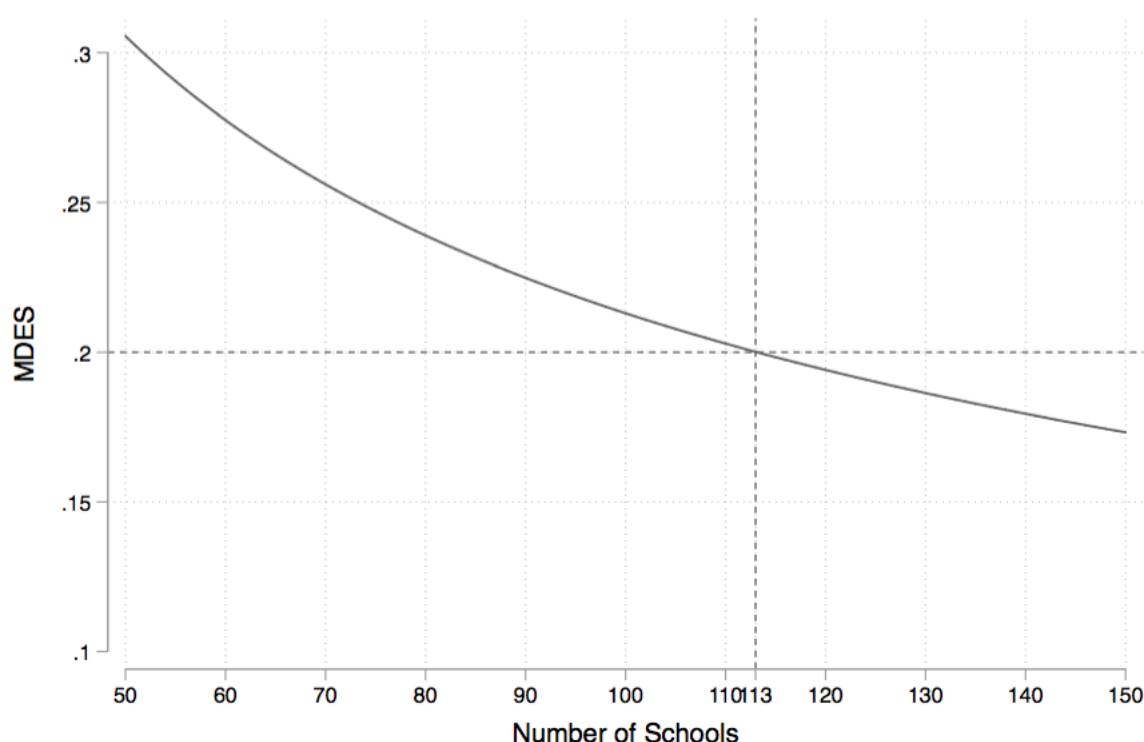
Sample size

Sample size calculations were carried out for the WAM, since this is the primary outcome of interest, and were carried out with an initial baseline of achieving a minimum detectable effect size (MDES), *d*, of 0.20 with the following assumptions: power of 0.8 for a two-tailed 0.05 significance test, randomisation at school level, an intra-cluster correlation (ICC) of 0.157 (EEF, 2015) and 25 pupils involved in the trial per school with 10 percent pupil-level attrition.

When estimating required effect sizes at the protocol stage, an appropriate pre-test/post-test correlation assumption could not be estimated empirically directly for this trial, since correlation data between the planned baseline and outcome measures were unavailable. This is because the planned baseline (score in the Year 1 phonics screening check, consistent with EEF policy to use an administrative measure rather than an additional pre-test) has only been in place since 2012, and our primary outcome measure (the WAM) is an even newer measure. EEF guidance suggests that a pre- and post-test correlation of 0.7 in education research is common (Torgerson & Torgerson, 2013); however, we see this as too optimistic in this case. The 21-day test–retest correlation coefficient of the WAM is reported to be 0.82 (Dunsmuir et al., 2015) but the time elapsed between the planned baseline and outcome measure in this trial would have been much longer, and it was never planned to use the WAM itself as a baseline. Our planned baseline measure (score in Year 1 phonics screening check) has less variance than would be ideal, due to a degree of bunching between the pass (32) and highest available mark (40). This bunching presumably occurs in order to push pupils over the pass line due to accountability concerns.

Nevertheless, given its closer temporal proximity to the outcome measure point, we expected (and continue to expect) that it is likely to explain more variance in our outcome than earlier measures available in the NPD (which would have to be measured at the Early Years Foundation Stage). While there is no direct measure of the correlation between the WAM and the phonics screening check available, a value of 0.52 was estimated, using Year 1 phonics screening check scores and Progress in International Reading Literacy Study (PIRLS) scores (DfE, 2017) (taken in Year 5, the same year as the WAM was administered). Given the similar time period between baseline and outcome test administration, and the related domain, we used this estimate as likely to approximate the value that would have been expected in this trial. Based on this, our sample size calculations at the protocol stage assumed that 25% of post-test variance at both pupil- and school-level would be explained by the pre-test (equivalent to pre-test/post-test correlation of 0.5).

---

[7] EEF guidance on ICCs (EEF, 2015) is provided for NPD outcomes. In the absence of ICC data for our outcomes of interest, we use this guidance specifically for the reading fine points score and, given uncertainty about the geographical spread of participating schools, we use the highest regional ICC (which happens to be Inner London) to the nearest two decimal places.

Figure 2: Minimum detectable effect size (MDES) estimate as a function of number of schools at design stage

These assumptions suggested a requirement of 113 schools to achieve an MDES of $d = 0.2$ (see Figure 2). Based on discussions between the evaluation team, the YJA team and the EEF at the set-up meetings, a target sample size of 100 was agreed. The YJA team confirmed that recruitment of 100 schools and intervention delivery to 50 treatment schools were reasonable and achievable numbers, given their capacity. Conditional on the sample size of 100 schools and the assumptions outlined above, this trial was estimated to be able to detect an effect of 0.21. Assuming the FSM sub-group is 15.3 percent of the total size of the sample (based on pupils aged 5–10 in data from DfE statistics (DfE, 2016) and ignoring that it may be higher if recruited schools are in more disadvantaged areas), and maintaining all other assumptions (which is likely to be a conservative approach, given lower levels of within-group variation in this sub-group), the MDES for this group at time of design was estimated to be approximately 0.32 standard deviations.

Ultimately, only 82 schools were recruited and included in the randomisation. As this was below the target level, albeit slightly offset by a small increase in the estimated average cluster size (from 23 to 24) in the schools recruited, the estimated primary analysis MDES increased to $d = 0.24$. However, there was a larger number of FSM pupils per school than expected (six pupils per school rather than three, after estimated attrition adjustment) and, as a result, the estimated MDES for this sub-group reduced a little to $d = 0.30$.

As noted above, there have been substantial changes to the analysis that we are ultimately able to carry out, due to data access issues with our planned baseline measures. These have a substantial bearing on the assumptions that underpin estimations of MDES. Specifically, the change in baseline variables included in the analysis model means that our early assumptions about the proportion of variance explained were optimistic (although, of course, this could still have been the case even if we had been able to obtain the planned baseline data): our data suggest that, with the substituted baseline measures, pupil-level variance explained (R^2) is 0.00 and school-level variance explained is 0.17 (with overall variance explained being 0.07). Furthermore, our average cluster size is lower than anticipated at 19, due to testing challenges; the number of control schools is 33 and the intervention schools is 39 (72 schools in total with a treatment ratio of 0.54; schools lost to follow-up are discussed below), and the intra-cluster correlation is substantially higher than anticipated (0.15 using standard EEF assumptions) at 0.29. Altogether, this leads to an estimated MDES for the primary analysis of d=0.36. This implies substantially reduced precision or, equivalently, more uncertainty (larger confidence intervals) around our treatment estimate than would have been the case if our initial assumptions were met – this higher level of uncertainty should be borne in mind in interpreting the findings.

There were similar challenges for the FSM sub-group. There, the pupil-level variance explained is again 0.00, the school-level variance explained is 0.17 (leading to an overall figure for variance explained of 0.04), the ICC is 0.26, and the average cluster size is three (the number of schools are as with the main sample exercise). This results in an estimated MDES for the FSM sub-group analysis of $d = 0.47$.

Randomisation

Schools recruited by the project team were randomly assigned by the evaluation team.

Stratification (referred to as blocking in the evaluation protocol) was used to attempt to improve cross-arm comparability of schools, and to improve precision of estimates. Eight strata were defined on the basis of:

- randomisation batch based on timing of sign-up;

- class composition by EAL (high vs. low split at sample median);

- class composition by eligibility for FSMs (high vs. low split at sample median).

This approach tried to ensure that our treatment and control groups is well balanced in terms of these characteristics, which are likely to be correlated with our outcome measures (EEF, 2018). High and low EAL and FSM in these definitions will be defined as above and below by the sample median (calculated separately for each randomisation batch) in each case, to ensure that block sizes are approximately equal (which may not be the case if we used population, rather than sample, characteristics). Table D7 in Appendix D shows the number of treatment and control schools in each stratum.

Randomisation followed recruitment of schools, including the signing of Memoranda of Understanding (MoU) and baseline data collection, in March–April 2018. This randomisation process was conducted using a script run in Stata 14. The randomisation followed the following process:

- The schools were stratified into four blocks on the basis of proportion of FSM students (split across the median sample proportion) and proportion of EAL students (split across the median sample proportion).

- Each school was assigned a randomly generated number (setting a stable seed for the random number generation).

- The schools were sorted by block and random number.

- Schools were assigned to the treatment arm and to the control arm in turn.

The code that was used to operationalise these steps (including the stable seed to allow for replication of the process) is reported in Appendix F of this report.

Statistical analysis

Primary analysis

Our primary analysis focused on the WAM score and was performed using Stata 15. All continuous variables were used in their 'raw' form (in line with EEF guidance), as there was no clear reason to transform the data.

An ordinary least squares (OLS) model was estimated, in which our outcome variable was regressed on a treatment arm indicator, strata indicators (based on proportion of the class eligible for FSM and proportion of the class identified as EAL), and in a deviation from the evaluation protocol, the following baseline variables: indicators of FSM eligibility, EAL status, class-level FSM composition and class-level EAL composition (further details below).

As noted by EEF guidance, in a model that does not account for clustering, when this is a feature introduced by the experimental design, 'the point estimates will be accurate, but the standard errors will be downward biased' (EEF,

2018, p.3). However, we accounted for the potential effects of the experimental design in this respect by calculating standard errors, taking into account clustering (Angrist & Pischke, 2008) at the school level, which allows for correlation of pupil outcomes within schools. We prefer this to the use of a hierarchical linear model that makes additional assumptions about the school-level effects that may not be justified. We also estimated randomisation–inference *p*-values accounting for the clustering and stratification of the design, which were consistent with those based on clustered sampling inference (see Appendix D).

The estimated impacts are intention-to-treat (ITT) effects, and have been reported with 95% confidence intervals (CIs). ICCs have also been reported.

In the evaluation protocol and SAP, we stated our intent to estimate the following model in order to estimate the ITT impact of the intervention:

$$Y_{ij} = \alpha + \beta_1 Treat_j + \beta_2 PreTest_{ij} + \gamma' X_j + \varepsilon_{ij},$$

where individual i is nested in school j, Y_{ij} is the WAM score, $PreTest_{ij}$ is the value of the phonics screening check score (using the NPD variable PHONICS_MARK) used as a pre-test, $Treat_j$ is our school-level treatment indicator, X_j is a vector of stratification variables, and ε_{ij} is an error term. Standard errors are calculated allowing for clustering at school level (j).

However, due to data access issues discussed above, we are unable to estimate this model due to the unavailability of PHONICS_MARK as $PreTest_{ij}$. Instead, we estimate the following model, in which $PreTest_{ij}$ has been replaced with FSM eligibility, EAL status, class-level FSM composition and class-level EAL composition
n (as discussed above):

$$Y_{ij} = \alpha + \beta_1 Treat_j + \beta_2 FSM_{ij} + \beta_3 EAL_{ij} + \beta_4 FSMProp_j + \beta_5 EALProp_j + \gamma' X_j + \varepsilon_{ij},$$

where everything is as per the planned model except that FSM_{ij} is whether individual i is eligible for FSM and, similarly, EAL_{ij} is whether individual i is recorded as having English as an Additional Language (EAL), while $FSMProp_j$ is the FSM composition of treated class in school j, and $EALProp_j$ is the same for its EAL composition.

As such, this report's primary intention-to-treat estimate is recovered from the estimate of β_1 in this latter model when it is estimated on the full sample at randomisation.

Note that while this model is a deviation from the evaluation protocol and SAP, it was planned and reported to the EEF ahead of analysis being carried out. The model has not been altered depending on the significance of any variables included (i.e. no variables were removed due to being statistically insignificant), including the vector of blocking variables (X_j). Syntax for this primary analysis model is reported in Appendix F.

Secondary analysis

We conducted two secondary outcome analyses:

Writing self-efficacy

- Same as the revised primary outcome analysis, except replacing Y_{ij} with the WSEM score. Note that this is a deviation from protocol, which stated that this would be the same as the planned primary outcome analysis except replacing Y_{ij} with the WSEM score and $PreTest_{ij}$ with assessment of pupils' Personal, Social and Emotional Development skills from the Early Years Foundation Stage Profile (aggregated scores from NPD FSP_PSE_G06, FSP_PSE_G07 and FSP_PSE_G08). This change was made due to data access problems rendering these NPD variables unavailable.

Ideation

- Same as the revised primary outcome analysis, except replacing Y_{ij} with the Ideation sub-score from the WSEM. Note that this is a deviation from protocol, which stated that this would be the same as the planned primary outcome analysis, except replacing Y_{ij} with the Ideation sub-score from the WSEM and $PreTest_{ij}$ with assessment of pupils' Personal, Social and Emotional Development skills from the Early Years Foundation Stage Profile (aggregated scores from NPD FSP_PSE_G06, FSP_PSE_G07 and FSP_PSE_G08). This change was made due to data access problems rendering these NPD variables unavailable.

It was also planned, potentially as part of a separate report, to estimate the impact on KS2 grammar, punctuation and spelling test attainment. Unfortunately, the relevant KS2 national curriculum tests in summer 2020 that would have collected these data were cancelled as a result of the COVID-19 pandemic. As such, this medium-term follow-up is no longer possible.

Analysis in the presence of non-compliance

The following criteria have been defined in the trial protocol as variables that can be used to assess compliance to the intervention. This draws principally on attendance data collected from the project team. Compliance was measured at the school level, which reflects the intervention delivery method. A school is considered to have complied if and only if the following two conditions were met:

1. schools must have allowed for all eight days to be conducted in schools with the YJA delivery team;

2. all schools must have uploaded at least 10 media items by the end of the intervention OR be considered by the school's mentor, in their professional judgement, to have adequately participated.

We used Complier Average Causal Effect (CACE) analysis (Gerber & Green, 2012) to estimate intervention effects on treated children. We estimated the CACE using two-stage least squares (2SLS) regression, by estimating a (first stage) model of compliance as follows:

$$Comply_j = \alpha + \beta_1 Treat_j + \gamma' X_j + \xi_{ij},$$

where $Comply_j$ is the binary compliance variable defined above, and ξ_{ij} is an error term. The predicted values of $Comply_j$ from the first stage are used in the estimation of a (structural) model of our outcome measure Y_{ij}. In other respects, the specification remains the same as the revised primary outcome ITT model. This second stage model is specified as follows:

$$Y_{ij} = \alpha + \beta_1 \hat{Comply_j} + \beta_2 PreTest_{ij} + \gamma' X_j + \omega_{ij},$$

where $\hat{Comply_j}$ are the predicted values of treatment receipt derived from the first stage model, and ω_{ij} is an error term. Our primary outcome of interest is β_1, which should recover the effect of the intervention among compliers.

We conducted this analysis using the 'ivregress' functionality of Stata to make necessary adjustments to standard errors (which have also been clustered at school level) due to the instrumental variables approach. We note the deviation to protocol due to these models being based on the revised primary outcome ITT model, rather than the planned primary outcome ITT model, which is for the same underlying reasons of data access.

Missing data analysis

We describe and summarise the extent of missing data in the primary and secondary outcomes, and in the model associated with the analysis. Reasons for missing data are also described.

For all models, we planned to implement a missing data strategy if more than 5% of data in the model was missing or if more than 10% of data for a single school was missing. The strategy would be followed separately for each instance of model and variable for which the threshold is exceeded.

We first assessed whether the missing data was missing at random (MAR), since this is a prerequisite for missing data modelling to produce meaningful results. To do this we created an indicator variable for each variable in the impact model, specifying whether the data was missing or not. We then used logistic regression to test whether this missing status could be predicted from the following variables: all variables in the analysis model plus eligibility for FSM (and proportion eligible for FSM in the school), and EAL status (and proportion EAL in the school). Where predictability was confirmed, we proceeded to the appropriate next step of this strategy.

For situations for which the MAR assumption appeared to hold and only the outcome variable in the model was missing, we re-estimated the treatment effect using our pre-specified model with the addition of the covariates found to be statistically significantly predictive of missingness of the outcome.

For situations for which the MAR assumption appeared to hold and any variable other than the outcome variable in the model was missing, we used all variables in the analysis model plus eligibility for FSM (and proportion eligible for FSM in the school), and EAL status (and proportion EAL in the school) to estimate a multiple imputation (MI) model using a fully conditional specification, implemented using Stata MI to create 20 imputed data sets. We re-estimated the treatment effect using each data set and took the average, and estimated standard error using Rubin's combination rules (Rubin, 2004).

Analysis that is altered following the missing data strategy (either on a multiply imputed data set or with additional variables) would only ever be viewed as a sensitivity analysis. As such, the main estimates of the effectiveness of the treatment are derived from complete case analysis only. However, the sensitivity of the estimates to missingness would be assessed by comparing this main analysis with those altered following the missing data strategy. For example, if the complete case analysis model were to imply effectiveness but the imputed estimate did not, we would assume that the missing data is missing systematically to such an extent as to invalidate the initial conclusion of effectiveness, which would then be stated in the reporting of the evaluation.

Sub-group analyses

Following EEF guidance, we first tested for an interaction of the treatment and FSM status. This was originally planned to be carried out using the NPD variable EVERFSM_6_P (in line with EEF guidance) and the following model:

$$Y_{ij} = \alpha + \beta_1 Treat_j + \beta_2 FSMEver_{ij} + \beta_3 Treat_j * FSMEver_{ij} + \beta_4 PreTest_{ij} + \gamma' X_j + \nu_{ij},$$

where individual i is nested in school j, Y_{ij} is the WAM score, $PreTest_{ij}$ is the value of the phonics screening check score used as a pre-test, $Treat_j$ is our school-level treatment indicator, $FSMEver_{ij}$ is an indicator of FSM eligibility (EVERFSM_6_P), $Treat_j * FSMEver_{ij}$ is an interaction between these two terms, X_j is a vector of stratification variables, and ν_{ij} is an error term.

However, in a deviation from protocol, due to unavailability of the NPD-derived indicator of FSM eligibility (EVERFSM_6_P) and the unavailability of the phonics screening check score (for data access reasons, as discussed above), this model was adapted as follows:

$$Y_{ij} = \alpha + \beta_1 Treat_j + \beta_2 FSM_{ij} + \beta_3 Treat_j * FSM_{ij} + \beta_4 EAL_{ij} + FSMProp_j + EALProp_j + \gamma' X_j + \nu_{ij},$$

where all terms are defined as per the planned FSM interaction model above or the planned primary analysis model. Standard errors have been calculated allowing for clustering at school-level (j). Syntax for this interaction model is reported in Appendix F.

If a significant interaction was found (i.e. the absolute value of the point estimate of β_3 divided by the school-level clustered standard error is greater than 1.96), then we would proceed to conduct a specific sub-group analysis for those who are identified by schools as eligible for FSMs ahead of randomisation using the same model as our revised primary analysis. We note that this is also a deviation from protocol, as we would have defined this sub-group as those who have ever been registered for FSMs in the NPD (identified using the variable EVERFSM_6_P) and used our planned primary analysis model.

But for the small difference in definition of FSM (which investigation with the DfE suggests is unlikely to be material as no cleaning of the data submitted by schools is carried out before it is made available in the NPD), this sub-group was

identified in the trial protocol and FSM pupils are a key sub-group to be analysed in all EEF trials. This FSM sub-group analysis was conducted for both the primary and secondary outcomes.

Additional analyses and robustness checks

No additional analyses were planned as part of the project's SAP. All additional analyses and robustness checks carried out should be considered exploratory only.

In addition to carrying out inference through school-level clustered standard errors, we also estimated randomisation inference *p*-values, in order to check the robustness of inference to this approach. As this was not planned in the SAP, this should be considered exploratory and will not be used to guide interpretation of the results. However, it will provide useful information about the extent to which there is variation between these different approaches to statistical inference.

We run three exploratory robustness check models based on potential issues identified in the course of analysis:

- Due to delays in the testing of some schools, we run a robustness check model in which we replicated the primary analysis model but added a control for the number days between the date that the first school was tested and the date that the school in question was tested. The logic for this is that delays to testing could have affected the dosage of the intervention.

- It is possible that variation in the approach of different markers (for example a degree of leniency by some markers, despite the steps documented above to maximise consistency) who marked the WAM (primary outcome measure) could affect the treatment estimate. While marking was blind to treatment assignment, the relatively small number of markers could lead to imbalance in such approaches by chance. As such, we run a robustness check model in which we replicated the primary analysis model but added marker fixed effects.

- Imbalance in school level KS1 scores between treatment and control group schools in the analysis sample is identified as part of our balance checks. To check whether such imbalance might explain our findings, we run a robustness check model which replicates the primary analysis model but includes average school level KS1 scores as an additional covariate. As it was not possible to link in average KS1 scores for all schools, multiple imputation was used for this analysis, carrying out 20 regression-based imputations of average KS1 score, using all covariates in the primary analysis model.

Attrition, and in particular imbalanced attrition between the treatment and control groups, is identified as part of our analysis. To check for bias from this, we carry out a multiple imputation replication of the primary analysis with 40 regression-based imputations of the primary outcome measure using all covariates in the primary analysis model, the treatment variable, and interactions between the treatment variable and our baseline covariates (FSM, EAL, and class FSM and EAL composition) to allow for potential differences in attrition mechanisms to the extent we can model these with available data.

Estimation of effect sizes

Hedges' *g* effect size was calculated as set out by Hedges (1981):

$$g = J(n_1 + n_2 + 2)\frac{\overline{x_1 - x_2}}{\hat{s*}},$$

where our conditional estimate of $\overline{x_1} - \overline{x_2}$ is recovered from β_1 in the primary ITT analysis model;

$\hat{s*}$ is estimated from the analysis sample as follows:

$$s^* = \sqrt{\frac{(n_1-1)s_1{}^2 + (n_2-1)s_2{}^2}{n_1 + n_2 - 2}},$$

where n_1 is the sample size in the control group, n_2 is the sample size in the treatment group, s_1 is the standard deviation of the control group, and s_2 is the standard deviation of the treatment group (all estimates of standard deviation used are unconditional, in line with the EEF's analysis guidance to maximise comparability with other trials);

and $J(n_1 + n_2 + 2)$ is calculated as follows:

$$J(n_1 + n_2 + 2) = \frac{\Gamma\left(\frac{n_1+n_2+2}{2}\right)}{\sqrt{\frac{n_1+n_2+2}{2}}\Gamma\left(\frac{n_1+n_2+2-1}{2}\right)},$$

where n_1 is the sample size in the control group and n_2 is the sample size in the treatment group, or if calculating $J(n_1 + n_2 + 2)$ proves computationally intractable[8] using the above method, we instead fall back on the following approximation:

$$J(n_1 + n_2 + 2) \approx (1 - \frac{3}{4(n_1+n_2)-9)}).$$

Ninety-five percent confidence intervals (95% CI) of the effect size have been estimated by inputting the upper and lower confidence limits of $\hat{\beta}_1$ from the regression model into the effect size formula.

Estimation of intra-cluster correlation (ICC)

In order to estimate the ICC of the outcome measure at school-level, we employed an empty variance components model, as follows:

$$Y_{ij} = \alpha + \eta_j + \varepsilon_{ij},$$

where individual i is nested in school j, Y_{ij} is the WAM score, η_j is a school-level random effect, and ν_{ij} is an individual-level error term. The school-level random effect is assumed to be normally distributed and uncorrelated with the individual-level errors. An empty variance components model is used to facilitate comparability between trials (and in line with EEF guidance).

The ICC itself was then estimated from this model using the following equation:

$$\rho = \frac{var(\eta_j)}{var(\eta_j)+var(\varepsilon_{ij})}.$$

In the SAP, we also intended to estimate the ICC of the planned baseline measure, the phonics screening check score. In a deviation from this plan for reasons of data access, as discussed above, this analysis was not conducted.

Longitudinal analysis

We had planned to estimate the effect of the intervention on pupil performance in KS2 national curriculum tests in English grammar, punctuation and spelling. Unfortunately, the relevant KS2 national curriculum tests in summer 2020 that would have collected these data were cancelled because of the COVID-19 pandemic, so this medium-term follow-up is no longer possible. It may ultimately be possible to consider analysing impacts on GCSE results (e.g. GCSE English, taken in 2025), however planning for this is beyond the scope of this report.

---

[8] The output of the gamma (γ) function in the Hedges' *g* correction factor (*J*) becomes large quickly, making this method of computation intractable where $n_1 + n_2$ is not small. As such, it can quickly become intractable. Thankfully, the approximate method tends towards the full correction factor quickly. As such, where the computational intractability is an issue, the approximate method is appropriate. In any event, the correction factor is likely to be small in this trial.

Implementation and process evaluation (IPE)

As part of the mixed method design of this evaluation, an IPE was conducted to complement the findings from the impact evaluation. The IPE involved members of the team with expertise and knowledge of the arts and education, which they fed into the design, conduct and analysis of the IPE. This section describes the IPE aims, data collection and analysis methods used.

The purpose of the IPE was to understand how YJA was delivered in schools and, in particular, to better understand the barriers and facilitators to implementation and delivering the intervention with fidelity. The IPE was also used to understand how the YJA intervention compared to usual teaching practice, and to determine the cost to schools of delivering the intervention. (See the Evaluation objectives section for the full list of research questions.)
Multiple sources of data were triangulated to best answer the IPE questions. The primary approach to IPE data collection consisted of collecting case study data from six schools in the intervention-arm of the study. Each case study involved the following methods: a semi-structured interview with the teacher whose class received the intervention; a semi-structured interview with a member of the SLT; informal discussions with pupils; and an observation of YJA teaching facilitated by the YJA mentor. A survey was also sent to all intervention and control group schools and administrative data were collected. In addition, observations were conducted at three schools between November 2018 and January 2019 to understand earlier intervention implementation. Intervention manuals and guidance were also reviewed to inform interviews and data analysis.

The research questions and the data collection methods used to address them are shown in Table 3.

Table 3: IPE methods overview

Data collection method (case studies)	Data analysis method	Participant groups	Target number of participants per case study	Actual number of participants /activities per case study	Total number of participants / activities	Research questions addressed
Semi-structured interviews	Framework approach	SLT	1	1	6	RQ1 (Fidelity), RQ2 (Mechanisms), RQ3 (Quality), RQ4 (Support), RQ7 (Engagement), RQ10 (Mechanisms)
	Framework approach	YJA mentor	1	1	6	RQ1 (Fidelity), RQ2 (Responsiveness), RQ3 (Quality), RQ4 (Support), RQ5 (Mechanisms), RQ6 (Quality), RQ7 (Engagement), RQ8 (Adaption/Quality), RQ10 (Mechanisms), RQ11 (Implementation)
	Framework approach	Teachers	1	1	6	RQ2 (Responsiveness), RQ3 (Quality), RQ4 (Support), RQ5 (Mechanisms), RQ6 (Quality), RQ9 (Reach), RQ10 (Mechanisms), RQ11 (Implementation)
Informal group interview	Framework approach	Pupils	3–5 pupils	Varied by school (5–15 pupils)	38	RQ8 (Adaption/Quality), RQ10 (Mechanisms), RQ11 (Implementation)
Observations	Framework approach	Pupils and teacher	1	1 YJA session	6 YJA sessions	RQ1 (Fidelity), RQ5 (Mechanisms), RQ6 (Quality), RQ7 (Engagement)
Observations	Framework approach	Training provider and teachers	1 whole school training day and 1 half-day training session	2 whole school training days (half-day at each); 1 half-day training session		RQ1 (Fidelity), RQ4 (Support),
Baseline survey	Descriptive statistics	Teachers	41 intervention	14 intervention		RQ1 (Fidelity), RQ4 (Support), RQ6 (Responsiveness), RQ9 (Reach)
Follow-up survey	Descriptive statistics	Teachers	41 control, 41 intervention	43 control, 40 intervention (prior to cleaning); 33 control, 35 intervention (post-cleaning)		RQ1 (Fidelity), RQ4 (Support), RQ6 (Quality), RQ9 (Reach)
Cost interview	Descriptive statistics	Teachers	4	4		

Sampling and recruitment

Case studies

Six case study schools were selected using a purposive sampling approach to capture the range of Year 5 classes that had received the YJA intervention during the 2018/2019 academic year. The primary sampling criteria were:

1. level of engagement in the intervention (defined as high where schools had uploaded 10 or more media items to the YJA website and low where schools had uploaded fewer than 10 items);

2. the proportion of pupils receiving FSM (defined as high or low based, on whether the school's FSM percentage was above or below the median for all YJA intervention schools, which was 13.7% [information obtained from UK Government, 2019]).

The secondary sampling criteria were:

1. geographical region (categorised as East of England, East Midlands, Greater London, North East England, North West England, South East England, South West England, West Midlands, and Yorkshire and the Humber);

2. Ofsted rating (recorded as Outstanding, Good or Requires improvement [information obtained from UK Government, 2019]).

Of the six schools initially sampled, one declined to participate, so a replacement school was sampled that had similar characteristics. Table 4 sets out the characteristics of the six case study schools that were recruited.

Table 4: Characteristics of case study schools

Characteristic	Number
Number of media items uploaded	
1–10	3
11–20	2
21–30	1
Free school meals (FSM) rate	
0–5%	1
6–10%	2
11–15%	0
16–20%	1
21–25%	1
26–30%	1
Geographical region	
East Midlands	1
Greater London	1
South East	1
North West	1
West Midlands	1
Yorkshire and the Humber	1
Ofsted rating	

Outstanding	1
Good	3
Requires improvement	2

Surveys

Baseline and follow-up surveys were administered to schools taking part in the trial. The baseline survey was distributed to schools randomised to the intervention-arm of the trial in December 2018. Unfortunately, the response rate (calculated using the number of schools originally randomised to the intervention-arm of the trial) was very low (34%), so it did not produce meaningful data. Instead, questions from the baseline survey were incorporated into the follow-up survey. For the follow-up survey, a census approach was taken, meaning that all schools involved in the trial were invited to respond. Depending on which arm of the trial they were randomised to, schools were sent either an intervention or control survey. Because we are not conducting any statistical inference with the survey results, confidence intervals are not given for response frequencies; in addition, as the schools taking part in the trial were not a representative sample of all primary schools in the UK, it is not appropriate to generalise the findings beyond this group of schools.

Data collection

Case studies

Sampled schools were contacted by email, and where schools agreed to take part, a date was arranged for a researcher to visit. The researchers did not inform the delivery team about which schools they intended to visit, although some schools informed their YJA mentor that a researcher would be visiting. All visits took place in April and May 2019 on a date when the YJA mentor was facilitating teaching at the school. At the visits, the following data were collected:

1. a semi-structured, audio recorded interviews with (i) the Year 5 teacher involved in delivery of the intervention; (ii) a member of the school's SLT; and (iii) the YJA mentor;

2. an observation of YJA teaching, facilitated by the YJA mentor;

3. informal discussions with participating children.

All of the instruments used for these purposes can be found in Appendices G–K. It was also set out in the protocol that interviews would be conducted with teachers both before and after the observation. The first interview was intended to focus on their broader experience of the programme; the second interview was intended to discuss the observation. However, when arranging interviews, teachers expressed concern at leaving their classroom twice, or had a timetable that did not support conducting interviews before and after instruction; instead it was agreed with teachers that they would attend one summative interview. Given the design of the interview schedule, this change did not pose any risk to the quality of the data being collected.

The interviews were conducted using guides that focused on exploring the following: usual practice in relation to teaching writing; the context in which the intervention was implemented; the facilitators and barriers to implementation; children's engagement; the perceived impact of the intervention; and the mechanisms underlying this change. Interviewees were informed that the interview was anonymous, that they could withdraw at any time, and that they did not have to answer any questions they did not want to. Discussions with children covered the content of YJA and how they found YJA activities; what they thought of the YJA mentor; and how YJA compares to other lessons. The researcher recorded notes during and immediately after each discussion. Full interview guides can be found in the Appendices.

Observations focused on the activities carried out as part of YJA; the role of the teacher and YJA mentor in delivery; children's engagement with the intervention; and the resources available to support delivery. The observations proforma was developed collaboratively in partnership with the UCL team, who brought their respective subject matter expertise. Specifically, Dominic Wyse has writing, music and mixed-methods research expertise; Gemma Moss has expertise in early literacy development; and Andrew Burn has expertise in English, media and drama. Observation notes were recorded first in field notes, then transferred to a structured proforma (which can be found in Appendix K). The observation data were used to help researchers probe effectively during the interviews and to deepen understanding of observed practice. Strategies to reduce bias were implemented, including having two researchers

(e.g. one BIT and one UCL team member) conducting the initial observations, discussing and agreeing upon the final observation notes that were recorded; one of those researchers continued to conduct the subsequent observations in the other case study schools.

Surveys

The research questions and programme logic model were used to inform the design of the baseline survey. Data from interviews, as well as feedback from the delivery team, were used to inform the design of the follow-up surveys. The survey administered to intervention schools covered the content of the eight-day YJA teaching programme; teachers' views on the YJA mentor; the class's engagement in YJA activities outside the eight-day teaching programme; children's engagement in YJA; and the perceived impact and quality of the intervention. The control survey focused on understanding usual practice in relation to teaching media and journalism. Full details on the intervention and control group surveys can be found in Appendices.

A link to the online survey platform SmartSurvey was sent to the key contacts at each school in May 2019. For the intervention survey, the email stated that the survey should be completed by the teacher who had been involved in the YJA intervention, while the control survey email stated that a Year 5 teacher should complete the survey. Reminders were sent to schools who did not initially complete the survey. Those schools that did not respond to reminders were followed up with a phone call and given the opportunity to complete the survey over the phone. All data were collected by the end of July 2019.

Administrative data

Data were collected on the number of YJA teaching days delivered in each school (out of a total of eight across the school year), and the number of media items uploaded by each school to the YJA website. These data were sent by the delivery partner to the evaluation team in March 2019 to inform case study sampling. As the number of teaching days varied minimally between schools, only the number of uploaded media items was used to inform case study sampling. At the end of the trial, the data was provided again, when a complete data set was available, as a measure of school engagement with the intervention.

Analysis

Case studies

Verbatim transcripts of the interviews and notes from the observations were analysed, using the Framework approach (Ritchie et al., 2013). Firstly, emerging themes were identified through familiarisation with the data. The analytical framework was then created using a series of matrices in Excel, each relating to an emergent theme. The columns in each matrix represented the key sub-themes drawn from the findings, and the rows represented individual participants interviewed or schools observed.

The interview and observation data were then summarised in the appropriate cell, which meant that all data relevant to a particular theme were noted, ordered and accessible, facilitating a systematic approach to analysis that was grounded in participants' and schools' accounts. Analysis involved working through the charted data to draw out the range of schools' experiences and participants' views, and identifying similarities, differences and links between them. Thematic analysis (undertaken by looking down the theme-based columns in the Framework) identified concepts and themes, and the case-based analysis (undertaken by comparing and contrasting rows in the Framework), enabled links within cases to be established and cases compared and contrasted with each other.

During the analytical process a balance was maintained between deduction (using existing knowledge and the research questions to guide the analysis) and induction (allowing concepts and ways of interpreting experience to emerge from the data). The *Implementation and process evaluation (IPE) findings* section is organised based on the identified themes and sub-themes, which are outlined and described in their respective subsections. Multiple strategies were employed by the researchers to increase the credibility (i.e. accurate representation of the data), transferability (i.e. potential to apply the findings to other settings), dependability (i.e. traceable, logical analytical process) and confirmability (i.e. being grounded and traceable to the raw data) of the findings, with the ultimate aim of reducing the bias during the analytical process (Hannes, 2011).

First, in terms of striving to increase credibility, the researchers conducted peer debriefing meetings with the senior qualitative research lead (Matt Barnard, Head of Evaluation at BIT) and qualitative researchers who were not directly

involved in the data collection or analysis process for the intervention. In addition, in accordance with the chosen approach to data analysis, the researchers focused on describing range and diversity, including the noting of any disconfirming cases. Verbatim participant quotations are used to provide evidence and exemplify the theme(s) discussed in the paragraph before the quotation. Quotations were selected by the qualitative researchers who conducted the data analysis, by considering multiple factors, including how well they exemplify the theme(s) discussed. The researchers also sought to ensure that the quotations used in the IPE findings capture the variation in terms of points of view and experiences, as well as type of participant interviewed (e.g. SLT, teachers, TAs) and the associated schools. Further, as qualitative data can only be generalised in terms of range and diversity and not in terms of prevalence, the analytical outputs focus on the nature of experiences, avoiding numerical summaries or language such as 'most' and 'majority'.

Second, to increase the potential for transferability and assessment of applicability to other contexts, the *Sampling and Recruitment* section describes the key details of the case study schools and the selection criteria. The findings also include important contextual details about the case study schools (e.g. motivations for choosing to implement the intervention).

Third, to increase dependability and confirmability of the findings, the researchers maintained a detailed audit trail and triangulated the data by comparing the findings from multiple types of participant (e.g. teachers, SLT) and sources of data (e.g. survey, interviews, observations). Researchers adhered to the key principles of the Framework approach, which includes ensuring that data management and analysis is systematic, comprehensive, transparent and grounded in the participants' accounts. Doing this was facilitated by the creation of a series of matrices in Excel that contained descriptive summaries of data that can be easily traced back to the verbatim quote on the relevant page of the transcript being described.

Surveys

Follow-up survey data were first cleaned by ensuring that all responses received came from a school in the relevant arm of the trial. Following this, data were checked to identify schools that had returned two or more survey responses. In these instances, the most recent survey completed by a respondent who identified as a 'teacher' was used for analysis. For instance, if there were two responses, both from a teacher, then the most recent one was kept, whereas if one response was from an SLT member and one was from a teacher, the SLT response was removed even if this was more recent. Data from teachers were prioritised where there were multiple responses from the same school, because it was assumed that the class teacher would have been more involved in the delivery of the intervention (for the intervention survey) and know more about usual classroom practice (for the control survey), and therefore their views and experiences were most relevant. Due to low response rate, baseline survey data were not analysed.

Prior to cleaning the survey data set, there were 40 intervention school responses and 43 control school responses. Following cleaning, there were 35 responses from intervention schools (out of 41 randomised to the intervention group) and 33 responses from control schools (out of 41 randomised to the control group), giving a response rate for intervention schools of 85% and for control schools of 80%. Stata (version 14) was used to generate descriptive statistics for each question. Percentage scores are reported, where relevant, in the *IPE findings* section. Complete survey findings are provided in the Supplemental Appendix.

Costs

The evaluation gathered three key categories of data: direct marginal costs (which will form the basis of the cost per pupil); pre-requisites (which will be reported separately from the cost per pupil); and school staff time.

The data was gathered in two ways. Firstly, the evaluators requested from delivery partners information on how much they charged schools for delivering the intervention as part of the evaluation, and how much they will charge schools in the future, excluding any funding or subsidy that is associated with delivering the intervention as part of this evaluation. The latter data are used in calculating the cost per pupil; the former data are to ensure there is clarity about the precise nature of the data that is being requested and transparency of the approach.

The second mode of data collection was the use of case study interviews, as specified in the protocol. IPE interviews were used to determine whether questions about costs would be included in surveys, with the decision taking into

account survey length and risk of damaging response rates. Based on low responses to the baseline survey, it was judged appropriate to omit cost-related questions so as to keep survey length down and not potentially dampen response rates with too many questions. Instead, costs were further explored through interviews, which were also deemed a more appropriate method to gather detailed data, as they allow for follow-up questions to clarify responses and probe for more information.

The evaluation team felt the programme cost was best estimated by having a good sense of the range and diversity of experiences, which was facilitated by using a case study approach supported by purposive sampling (Ritchie et al., 2013). We selected case study schools from those who had good engagement with the programme, as they were more likely to give the best indication of the resources needed to implement the programme fully; including schools with little engagement was likely to artificially deflate costs. As an indicator of this, we selected schools from the pool who had completed outcome data collection (as fidelity data was not available at that stage).

The resources required to deliver the intervention were most influenced by staff time and any related marginal costs (such as travel and subsistence). The evaluators assumed these things were most likely to be related to the amount a school has to spend per pupil and the nature of the local area and school population, the most relevant indicator for which is the percentage of pupils eligible for FSM. That is, school spending on the programme was likely related to school financial resources. Therefore, we purposively selected one school in each of the following four categories:

1.   School with percentage of FSM in top half of participating schools and spend per pupil in top half;

2.   School with percentage of FSM in top half of participating schools and spend per pupil in bottom half;

3.   School with percentage of FSM in bottom half of participating schools and spend per pupil in top half;

4.   School with percentage of FSM in bottom half of participating schools and spend per pupil in bottom half.

The case study data were collected by research assistants (RAs) employed and trained by BIT. The RAs conducted interviews with teachers via telephone using a structured interview guide designed by BIT for this purpose. RAs estimated it took approximately 20 minutes to complete the discussion.

Teachers were asked to report on direct costs of the intervention to the school, materials purchased, travel and subsistence, the cost of covering staff at training and the cost of any new physical materials purchased to improve the classroom environment. Teachers also reported on time spent embedding the intervention in their school, time at training, as well as time spent preparing to deliver the intervention. Staff were also asked to report on time taken to organise supply cover and the amount of supply cover.

Data from these interviews were used to calculate the financial and time costs outlined in this report.

Timeline

Table 5: Timeline (a) Overall evaluation timeline

Dates	Activity	Staff responsible / leading
June 2015	Participants sit Year 1 Phonics Screening Check, planned as baseline measure	N/A
November 2017–April 2018	Recruitment: The YJA team begins recruitment halfway through the Autumn term 2017 using its existing network of schools and continue into early 2018	YJA team
November 2017–February 2018	Pre-randomisation data collection	BIT and YJA team
17 September 2018	First randomisation batch	UCL
8 October 2018	Second randomisation batch	UCL
September 2018–June 2019	Intervention in schools	YJA team

September 2018–June 2019	Implementation and Process Evaluation fieldwork (see further details on IPE timeline below)	BIT and UCL
May–July 2019	Outcome testing: Pupils' writing and self-efficacy outcomes will be collected by BIT. These tasks marked by PGCE students at UCL in a process overseen by BIT.	BIT and UCL
July–September 2019	Collation and cleaning of outcomes and compliance data in readiness for upload to ONS SRS for linkage with DfE's NPD extract	UCL and BIT
January–July 2020	Project paused awaiting conclusion of data sharing agreement necessary for upload of project data to the ONS SRS for linkage with NPD.	
August–October 2020	Project resumed with planned revisions as a result of delays in achieving data sharing agreement. Impact analysis and report writing. UCL led on the data analysis with agreed deviations from published statistical analysis plan (SAP).	UCL and BIT

(b) IPE timeline

Date	Activity
Autumn term 2018	Observation of training of whole school day visit
	Collection of baseline survey
	Collection of school characteristics
Spring term 2019	Observation of mid-point YJA half-day school session
	Collection of fidelity data to inform case study sampling
	Finalised sampling strategy
	Conducted in-school case studies
Summer term 2019	Conducted in-school case studies
	Administered follow-up survey
Autumn 2019	Conducted analysis

35

Impact evaluation results

Summary

- There was no significant impact of participating in the YJA on writing attainment (WAM);

- There was no significant impact of participating in the YJA on writing self-efficacy (WSEM) or on ideation;

- There was no differential effect of the intervention for FSM eligible pupils.

Participant flow including losses and exclusions

The flow of participants is detailed in Figure 3. Of the original 3,500 schools that were approached, 3,247 did not respond, zero did not meet the inclusion criteria, and 172 declined to participate for other reasons (e.g. did not return the MoU on time). In total, 82 schools agreed to participate in the trial and met the eligibility criteria. These schools were randomly allocated to the intervention and control groups using a block stratified randomisation, as described above. At randomisation, 1,072 students in 41 schools were allocated to the intervention group and 1,065 students in 41 schools to the control group.[9]

One treatment and eight control schools declined to participate in testing at the end of the intervention, and some pupils were absent at the point of testing (discussed further below), such that outcome data was not collected for 185 students in the treatment group and 339 students in the control group. As is the case in many trials, it was easier to maintain contact with treatment schools and the developers were able to help ensure that fewer treatment schools dropped out than control schools. This meant that, ultimately, 887 pupils in 40 schools allocated to treatment and 726 pupils in 33 schools allocated to control were analysed. We recognise potential concerns from this imbalance in attrition (although we also note recent evidence that attrition from RCTs causes less bias than commonly assumed, particularly when from the control group, see Weidmann & Miratrix, 2021) and return to this point in the additional analysis below.

The MDES estimated at various points of the trial is reported in Table 6. This was 0.21 at the design stage, increased to 0.24 at randomisation (primarily due to under-recruitment), and increased again to 0.36 at the stage of analysis. This increase may be attributed to a combination of being unable to access the planned baseline measures (discussed further above), attrition rates, and a higher ICC of the outcome measure than was anticipated based on previous EEF guidance.

---

[9] These numbers differ slightly from those reported in the SAP (2,157 total pupils, 1,091 in intervention schools and 1,066 pupils in control schools). The difference in intervention pupils is due to a school providing an entire year group list instead of a specified class list. The difference in the control pupils is a typo in the SAP.

Figure 3: Participant flow diagram

Recruitment

Approached (school *n* = 3500)

Did not agree to participate

Agreed to participate (school *n* = 254)

Excluded (school n = 0)

Not meeting inclusion criteria (school *n* = 0)

Other – MoU not returned on time (school *n* = 172)

Allocation

Randomised

Treatment

(school *n* = 41; pupil *n* = 1,072)

Control

(school *n* = 41; pupil *n* = 1,065)

Follow-up

Lost to follow-up

(school *n* = 1; pupil *n* = 185)

Post-test data collected

(school *n* = 40; pupil *n* = 887)

Lost to follow-up

(school *n* = 8; pupil *n* = 339)

Post-test data collected

(school *n* = 33; pupil *n* = 726)

Analysis

Not analysed

(school *n* = 0; pupil *n* = 0)

Analysed (school *n* = 40; pupil *n* = 887)

Not analysed

(school *n* = 0; pupil *n* = 0)

Analysed (school *n* = 33; pupil *n* = 726)

Table 6: Minimum detectable effect size (MDES) at different stages

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
MDES		0.21	0.32	0.24	0.30	0.36	0.47
Pre-test/post-test correlations	Level 1 (pupil)	0.50	0.50	0.50	0.50	0.00	0.00
	Level 2 (class)	0.50	0.50	0.50	0.50	0.17	0.14
	Level 3 (school)	N/A	N/A	N/A	N/A	N/A	N/A
Intra-cluster correlations (ICCs)	Level 2 (class)	0.15	0.15	0.15	0.15	0.29	0.26
	Level 3 (school)	N/A	N/A	N/A	N/A	N/A	N/A
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?		two-sided	two-sided	two-sided	two-sided	two-sided	two-sided
Average cluster size		23	3	24	6	19	3
Number of schools	Intervention	50	50	41	41	40	40
	Control	50	50	41	41	33	33
	Total:	100	100	82	82	73	73
Number of pupils	Intervention	1150	150	1072	298	887	244
	Control	1150	150	1065	265	726	169
	Total:	2300	300	2137	563	1613	413

Notes. Please note that there is a correction to the numbers of pupils reported under the randomisation phase, compared to the SAP. This is due to (a) over-counting of the number of pupils considered to have been randomised due to two schools initially submitting more pupils than a single class and this not being able to be corrected ahead of randomisation itself; (b) a typographical error which increased the number of control group pupils reported in the SAP by one. As no pre-test data was ultimately available, the pre-test/post-test correlations as analysed report the multiple correlation between the covariates included in the analysis model and the primary outcome measure. Average cluster sizes are harmonic means of cluster sizes which is more conservative than using the arithmetic mean in the presence of unequal cluster sizes (Bulus et al., 2019).

Attrition

One treatment and eight control schools declined to participate in testing at the end of the intervention, and a further 155 pupils in treatment schools and 120 pupils in control schools were absent at the point of testing (despite return 'mop-up' testing in schools where more than a small number of pupils were not available on the day of main testing). Ultimately, 887 pupils in 40 schools allocated to treatment and 726 pupils in 33 schools allocated to control are analysed. This led to an attrition rate of 17.3 percent for the intervention group and 31.8 percent for the control group, which amounted to 24.5 percent of the total randomised sample (see Table 7).

Table 7: Pupil-level attrition from the trial (primary outcome)

		Intervention	Control	Total
Number of pupils	Randomised	1,072	1,065	2,137
	Analysed	887	726	1,613
Pupil attrition (from randomisation to analysis)	Number	185	339	524
	Percentage	17.3%	31.8%	24.5%

Pupil and school characteristics

Table 8 presents the baseline characteristics of treatment and control schools and pupils as randomised. In general, it shows that treatment and control schools are broadly similar to each other and similar to the national average for a range of characteristics. There are, however, some differences.

Schools randomised to the control group were nearly half as likely as intervention schools to be rated 'Outstanding' by Ofsted (11 percent of control schools vs. 20 percent of intervention schools). This was also lower than the national average of 17 percent of schools being rated 'Outstanding'. Control schools were also more likely to be academies (47 percent) than treatment schools (37 percent), both of which were higher than the national proportion of 24 percent. We note, however, that this imbalance is slightly less pronounced in the primary analysis sample (see Table D1 in Appendix D).

Intervention schools had higher average KS1 performance (16.22 for intervention and 15.95 for treatment schools). This is a standardised difference of –0.233 (Imbens & Rubin, 2015). This imbalance could potentially bias the effect of the YJA intervention in a positive direction, since intervention schools were already higher performing before receiving any treatment. We note, however, that this imbalance is less marked in the primary analysis sample (see Table D1 in Appendix D) where the analogous difference is –0.107.

Differences in pupil characteristics between the two groups of schools were smaller. Intervention schools had a slightly higher proportion of 'ever FSM' pupils (27 percent) vs. control schools (24 percent), which was also lower than the national average (31 percent). The standardised difference between treatment and control groups for this measure is –0.069 (Imbens & Rubin, 2015). The treatment and control pupils were somewhat more likely to have EAL (21 percent of intervention pupils and 24 percent of control pupils) as compared to only 15 percent at the national average. Here, the absolute standardised difference between the intervention and control groups is –0.063. Table D1 in Appendix D presents the analogous balance characteristics for the groups as analysed.

Table 8: Baseline characteristics of groups as randomised

School-level (categorical)	National-level percentage	Intervention group		Control group		
		N (missing)	%	N (missing)	%	
Urban	87.3	31 (7)	89.6	34 (1)	88.1	
Rural	12.7	3 (7)	10.4	6 (1)	11.9	
Ofsted: Outstanding	17.1	8 (3)	20.2	4 (1)	10.9	
Ofsted: Good	69.4	26 (3)	70.3	30 (1)	76.5	

		N (missing)	Mean (SD)	N (missing)	Mean (SD)	
Ofsted: RI/Inadequate	13.4	4 (3)	9.6	6 (1)	12.6	
School type: Academy	23.6	14 (2)	37.2	19 (0)	47.1	
School type: Community	41.2	14 (2)	34.2	10 (0)	25.7	
School type: Other	35.2	11 (2)	28.6	12 (0)	27.1	
School-level (continuous)	National-level mean	N (missing)	Mean (SD)	N (missing)	Mean (SD)	Standardised difference
KS1 average performance	15.9	36 (5)	16.22 (1.27)	38 (3)	15.95 (1.04)	0.233
Pupil-level (categorical)	National-level percentage	n (missing)	%	n (missing)	%	
FSM	30.9	291 (13)	24.2	258 (0)	27.1	
Non-FSM	69.1	781 (13)	75.8	807 (0)	72.9	
EAL	15.3	224 (42)	24.2	258 (0)	21.5	
Non-EAL	84.7	819 (42)	75.8	807 (0)	78.5	

Note. School-level imbalance is calculated applying weights for the size of the school, so that schools that are relatively more important in the pupil-level impact estimation are afforded the same importance in understanding imbalance.

Outcomes and analysis

Primary analysis

Table 9 presents the results of the intention to treat analysis for the primary outcome measure. It shows the unadjusted mean for the YJA intervention group (17.03) and the unadjusted mean for the control group (17.90). After adjusting for covariates in the analysis model, we find an adjusted mean difference of –0.66. Based on this, we calculate a Hedges' g effect size of –0.13, which is negative, and of a medium magnitude (equivalent to falling behind two months). It was not found to be a statistically significant difference between the two groups ($p = 0.16$).

The same unadjusted WAM scores for the intervention and the control group can be seen in Figure 4. The overall mean for the WAM across both treatment arms is 17.42 and the median is 18, with the similarity between the mean and the median reflected in the normal distribution of this outcome measure.

Conditioning the WAM scores on the variables used for stratification does not significantly alter the distribution for the treatment and control arms of this trial. The small mean difference between the two arms in our main analysis is evident in the plot: the solid treatment line is below the dashed control line.

Table 9: Impact estimates

| | Unadjusted means | | | | Effect size | | |
| Outcome | Intervention group | | Control group | | | | |
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges' g (95% CI)	p-value
Primary outcome							
WAM score (ideas scale double weighted)	915 (170)	17.03 (16.45, 17.62)	726 (339)	17.90 (17.07, 18.73)	1,613 (887; 726)	−0.13 (−0.32, 0.05)	0.16
Secondary outcomes							
WSEM Score	889 (196)	63.88 (62.84, 64.92)	709 (356)	64.08 (62.88, 65.28)	1,571 (862; 709)	0.03 (−0.09, 0.16)	0.61
Ideation score	889 (196)	19.85 (19.54, 20.15)	709 (356)	19.91 (19.49, 20.33)	1,571 (862; 709)	0.03 (−0.10, 0.16)	0.68

Figure 4: Histogram of unadjusted WAM scores by treatment arm
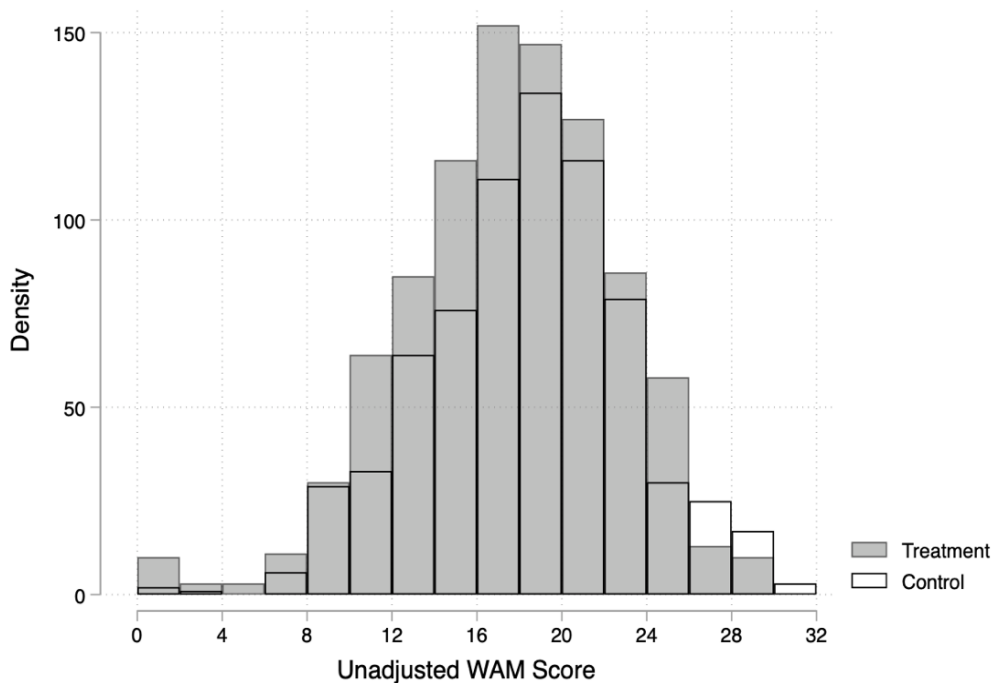
We also present the impact analysis results for the primary outcome measure graphically, through the second kernel density plot by treatment arm displayed in Figure 5. The adjusted WAM scores for this plot are obtained from regressing the WAM writing score on pre-test measures (planned to be phonics score but, due to data access issues, instead FSM and EAL status plus class composition of these) and the variables used for stratification (i.e. the analysis model other than the treatment indicator).

Figure 5: Histogram of adjusted WAM scores by treatment arm
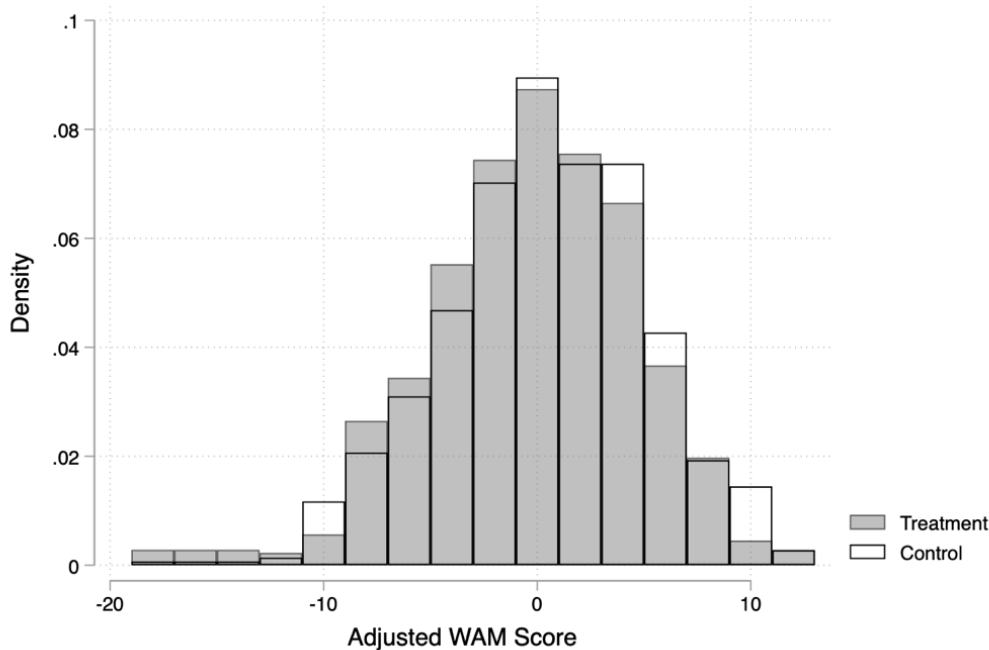
Secondary analysis

The results of the secondary analysis are also presented in Table 9. The secondary outcome measures for this trial are writing self-efficacy (WSEM) and ideation. These outcomes are more closely aligned with the content and aims of the YJA intervention and so one might expect them to be more likely to register a larger effect.

For the WSEM, there is a difference in the unadjusted means between the intervention group (63.88) and the control group (64.08), but it is small in magnitude. Figure 6 shows this very small difference in the WSEM between the two groups graphically. It shows a right-skewed distribution, indicating a high proportion of pupils giving high responses on how they view themselves as writers, introducing some risk of ceiling effects in the analysis of this outcome measure, which could attenuate the impact estimate.

Figure 6: Histogram of unadjusted WSEM scores by treatment arm
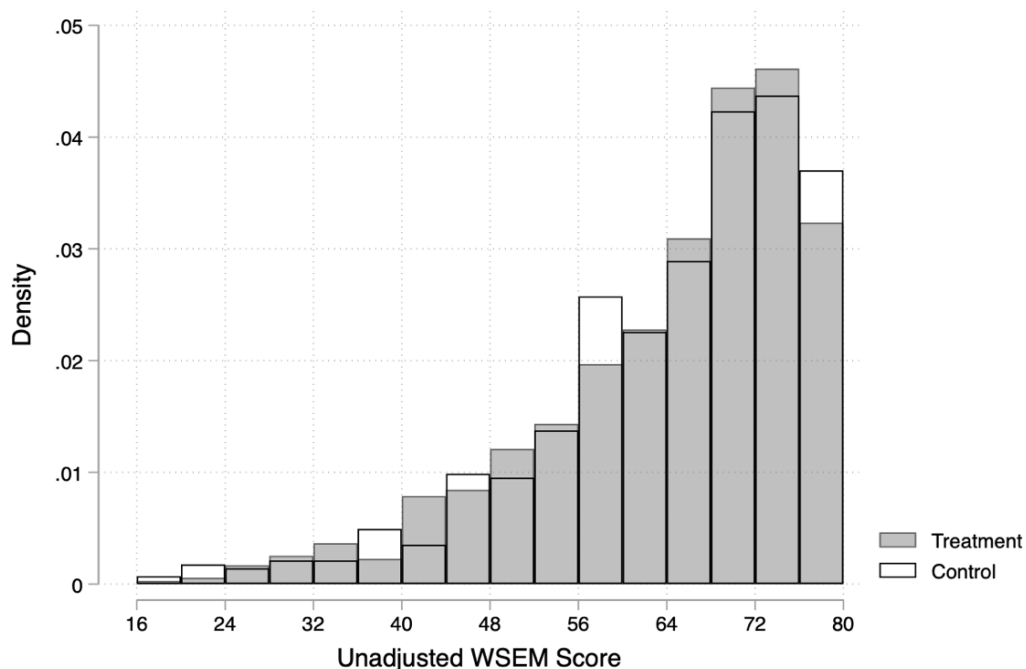
After adjusting for pre-intervention covariates the adjusted difference in means is 0.39, which translates into an effect size of 0.03. Note that this effect is no longer negative when we condition on covariates. This effect size is not statistically significant at the five percent significance level (*p*-value is 0.61). This may be seen in Figure 7, which shows the adjusted kernel density plots of the WSEM scores by treatment arm. There is no substantial difference between the solid and the filled bars in this figure.

Figure 7: Histogram of adjusted WSEM scores by treatment arm
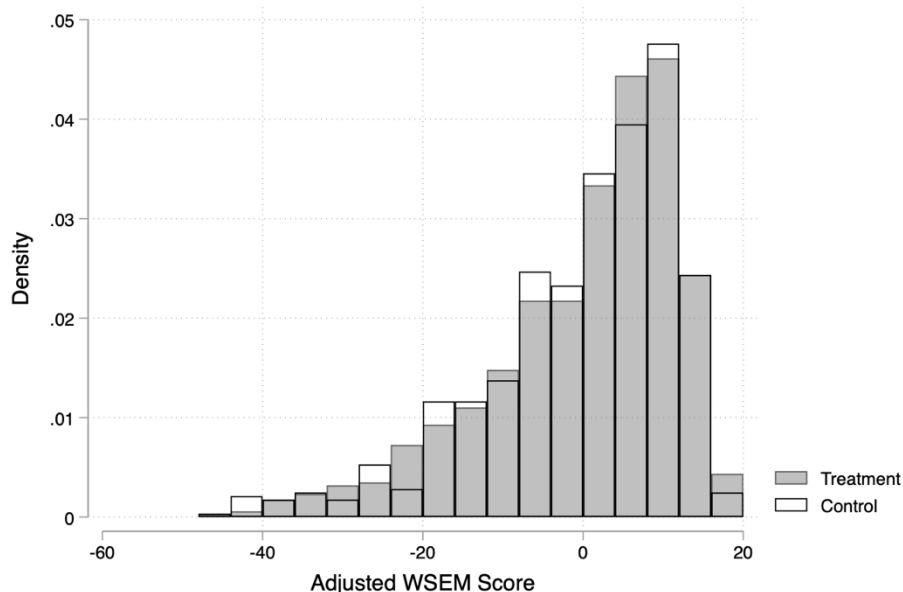
A similar result emerges for the ideation score. The control group has a higher mean ideation score (19.91 for the control group and 19.85 for the intervention group, see Table 9 for further details and Figures E4 and E5 in Appendix E), which yields an effect size of 0.03. This effect size is small in magnitude and not statistically significant (*p*-value is 0.68).

Analysis in the presence of non-compliance

In order to examine the issue of non-compliance, we estimate the Complier Average Causal Effect (CACE). As outlined in Figure 3, there was only one intervention school that did not comply, based on the compliance criteria outlined in this report. Nevertheless, we want to explore whether non-compliant implementation may be diluting an underlying treatment effect in compliant schools, so we account for this lack of fidelity in calculating the effect size for the primary outcome measure.

To calculate the CACE, we use the 'ivregress' functionality of Stata to make necessary adjustments to standard errors (which are also clustered at school level) due to the instrumental variables approach. Further details of this approach are available in the syntax file in Appendix F. As compliance data are available for the full primary analysis sample, this analysis is also carried out for this full sample (*n* = 1,613).

The extremely high correlation between treatment status and compliance status (*r* = 0.98) suggests that there is unlikely to be a difference in findings between the ITT and CACE analyses, while the large F test of the first stage of the instrumental variables model ($F_{(12, 71)} = 2494$) implies that there is no weak instrument issue, so the analysis is meaningful. The results of the complier analysis produce an effect size of –0.14 (95% CI: –0.33, 0.05), which is qualitatively similar to that of the ITT analysis; furthermore, the *p*-value of the CACE treatment estimate (0.15) shows that this effect size is not statistically significant. As such, we interpret these findings as not providing evidence of differential treatment effects among schools with higher levels of compliance for the primary outcome measure.

Missing data analysis

In the SAP for this trial, we outlined a missing data strategy. We noted that this strategy would be implemented if more than 5% of data in the model is missing or if more than 10% of data for a single school is missing. Due to issues with testing, a substantial number of pupils in treatment and control schools do not have outcome data. In order

to test whether or not this data is missing at random, we ran a logistic regression model to predict missingness in outcome data, using all variables in the analysis model plus eligibility for FSM (and proportion eligible for FSM in the school), and EAL status (and proportion EAL in the school).

The results of this analysis show that only the proportion of FSM pupils in the school is a statistically significant predictor of missing outcome data. This implies the potential for bias to complete case analysis ignoring this factor since, for example, FSM pupils might have worse outcome data, which we would be more likely not to observe due to missingness. However, since FSM is already included in the (revised) primary analysis model, there are no further actions possible in relation to systematic predictors of missing outcome data and this should address concerns about systematic missing outcome data associated with this factor (i.e. we believe, based on our analysis, that conditional on these variables outcome data are MAR). If any additional predictors of missingness had been discovered in the logistic regression modelling of outcome data missingness, we would have run an additional ITT model controlling for them, as well as a robustness check.

In terms of missing predictor data, none of our sample is missing these data (0.0% for EAL and 0.0% for FSM). Since no one in our sample is missing this information, there is no reason to undertake multiple imputation.

Sub-group analyses

As is standard in all EEF funded evaluations, we considered whether there is evidence of differential effects among pupils eligible for FSM as a separate sub-group ($n = 413$). We started by considering an augmented version of the primary analysis model including an interaction term between the treatment variable and membership of the FSM sub-group. The estimate of this interaction is small and not statistically significant ($p = 0.56$), providing little evidence of a differential effect among the FSM sub-group. (See Table D2 in Appendix D for full details.)

Additional analyses and robustness checks

As outlined in the SAP for this trial, no additional analyses were planned for this trial. However, in light of ongoing discussions about appropriate interpretation of classical statistical inference, we carried out alternative statistical inference, using randomisation inference to provide useful information on the extent to which there is variation between these different approaches to statistical inference. Randomisation inference is a method of conducting statistical inference using the uncertainty inherent in the randomisation process regarding the assignment of units in the trial to the treatment arms, rather than any appeal to an external sample and sampling variation (i.e. focusing on internal rather than external validity; see Cunningham, 2021). These comparative findings are reported in Table D4 in Appendix D. They result in very similar *p*-values, despite their differing conceptual underpinnings.

Furthermore, we decided to run four exploratory robustness checks to probe the main findings. The results of these analyses are reported in Table D6 in Appendix D.

Due to delays in testing some schools, we ran an additional robustness check, controlling for the date of testing since this could have affected the dosage of the intervention. The results obtained from this analysis were not substantially different from the overall impact evaluation results.

We also ran a model accounting for differences in the markers who marked the WAM (primary outcome measure), since it is possible that they introduced a new source of bias due to different degrees of leniency in marking. The inclusion of these marker fixed effects attenuated the effect size towards zero, but the effect size is still small in magnitude and negative.

Due to the imbalance in school-level KS1 scores, we also ran a model which included average school-level KS1 scores. Again, this analysis also did not substantially affect the overall effect size calculated. Imbalance in school-level Ofsted ratings was also discussed above. As such, it would seem a natural extension to run an analogous model including this as a covariate in our analysis model. Unfortunately, due to COVID-19 restrictions, we are unable to link this characteristic to our pupil-level analysis data and are not able to undertake such an analysis. However, we note relevant points that may help to provide some guidance on what we would expect from such an analysis. Ofsted ratings are generally stronger among schools in the intervention group and, as such, to the extent that we might expect a positive correlation between Ofsted ratings and pupil attainment, we would anticipate this to, if anything, increase the magnitude of the negative estimated treatment effect. However, we know that Ofsted rating is not a strong

predictor of pupil attainment (von Stumm et al., 2021), particularly when one has already conditioned on prior attainment, as we have done in the analysis with KS1 scores above. We would therefore not expect this to affect our treatment estimate dramatically.

Finally, imbalance in attrition has also been highlighted as a potential concern in this trial. In our final robustness check, we carry out a multiple imputation replication of the primary analysis with 40 regression-based imputations of the primary outcome measure, using all covariates in the primary analysis model, the treatment variable, and interactions between the treatment variable and our baseline covariates (FSM, EAL, and class FSM and EAL composition) to allow for potential differences in attrition mechanisms to the extent we can model these with available data. This analysis did not reveal any evidence of bias.

Estimation of effect sizes

As previously outlined, effect sizes are calculated using Hedges 'g. These have been presented in Table 9 for the primary and secondary outcomes. Appendix Table C1 in Appendix C contains additional information used in the estimation of effect sizes (e.g. the standard deviations).

Estimation of ICC

As this is a clustered trial, we estimate the ICC at the class level. At the time of randomisation, the ICC was estimated to be 0.15, which we based on EEF guidance drawing on analysis of NPD data. However, the ICC of the primary outcome measure (the WAM) has turned out to be considerably higher, at 0.36. There are aspects of the sample which may explain some or all of this difference, including that the schools participating in the trial are disproportionately drawn from urban areas (although this was also the case for the estimate using EEF guidance, since we used the highest regional ICC, which happened to be for London), and that our missing data analysis suggests the potential for our sample to differ in terms of FSM composition. That said, we doubt these explain the extent of the difference, which is likely attributable to the outcome and/or outcome measure itself, for example reflecting within-class shared writing practices that are consistently captured by the measure. This should be borne in mind in the design of future trials making use of measures of this type.

Implementation and process evaluation results

Summary

- Usual practice in the majority of intervention and control schools already included elements of journalism (e.g. writing newspaper articles).

- Some teachers found it difficult to reconcile the amount of time required for the YJA intervention with teaching the school curriculum.

- More than half of surveyed teachers (64%) reported they had engaged in YJA activities more than once per month (outside the eight teaching days). SLT support was seen as a facilitator for embedding the programme outside of the scheduled teaching days.

- Some teachers reported a lack of access to technology as a barrier to effectively implementing the YJA intervention.

- Teachers highlighted pupil enjoyment and engagement with the intervention, but felt that the full-day length of each training session was too long for some pupils.

- Teachers perceived the programme to have a positive impact on pupils. Of the teachers surveyed, 69% thought that YJA had a positive impact on pupils' writing (with 26% thinking it had a very positive impact). However, some teachers were uncertain about whether YJA would have improved writing attainment, though these teachers said the programme may have had more of an impact on engaging more reluctant writers and increasing writing confidence.

This section of the report contains the findings of the IPE. The first subsection contains background context on the case study schools and information on usual practice. The second subsection discusses factors affecting the fidelity of implementation of the YJA programme. The final subsection covers the mechanisms underpinning the effectiveness of the intervention and its perceived impact. Further details on where research questions are addressed is in Table 10.

Table 10: Findings sections and research questions

Section in IPE findings	Research question (RQ)
Context	RQ1. What other support do the pupils access to support their communication in both control and treatment?
	RQ3. How was the quality of the intervention perceived by teachers, senior leaders and TAs?
Implementation	RQ1. In what ways was the programme implemented? What are the barriers and facilitators of delivery? In particular: SLT 'buy-in' delivery of the intervention: (a) whether it appears to facilitate children's engagement
	RQ2.To what extent did the schools engage with the intervention in line with the intervention aims?
	RQ4.To what extent is the knowledge of arts practitioners and other practitioners integrated with the knowledge of teachers involved?
	RQ6. The relationship between the 'core' team and the rest of the class. How does this affect engagement with the programme? (Responsiveness) [Deviation: How did the core team selection process work (barriers and facilitators) and how did ways of working that affect the team's ability to function? (Quality)]
	RQ7. How do schools engage more broadly in showcasing journalist outputs? (Engagement)
	RQ8. Which elements of the intervention are most widely adopted and how does this affect outcomes? (Adaption/Quality)
	RQ9. To what extent is the intervention disseminated across the school? (Reach)
	RQ11. To what extent do school facilities affect the intervention?
Perceived mechanisms and impact	RQ5. What are the mechanisms that are taking place in the intervention, and to what extent are they bringing about change?
	RQ10. How does the YJA intervention affect literacy in the class and school? (Mechanisms)
	RQ3. How was the quality of the intervention perceived by teachers, senior leaders and TAs?

Context

This section discusses two contextual factors in case study schools:

1.　　the school's FSM rate;

2.　　the teacher's experience of journalism, which influenced the implementation of YJA.

Each is discussed in turn below, followed by a description of how YJA compares to usual practice.

The influence of the proportion of FSM-eligible pupils

Respondents from case study schools with a high proportion of FSM pupils tended to report also having a high proportion of pupils who spoke English as an additional language (EAL). This was a key issue raised by teachers, as it was felt to be more difficult to initially engage these pupils with YJA and to teach writing-related tasks. Teachers felt that this was due to the fact that these pupils sometimes struggled to understand English grammar and idioms, or had not been introduced to the vocabulary to describe certain objects.

> *'So just little things like that and you know, names of some animals and things like that, that the children just haven't heard of. Obviously, they might have a different word in their language, so they just don't know the English vocabulary for it.'* (Teacher 01)

Some teachers also highlighted the fact that these pupils could lack support to embed their learning from YJA at home. It was reported that cultural and religious commitments meant that some pupils had limited time to complete YJA-related activities outside school. Furthermore, some teachers said where parents spoke limited English, it was difficult for them to support their children with activities such as article writing.

Teachers in schools with a high rate of FSM also reported that they had a higher proportion of pupils with special educational needs (SEN). Teachers commented that these pupils generally found it harder than other pupils to engage in the YJA programme, and struggled with some tasks. For instance, students with dyslexia were reported to have particular difficulty with writing tasks and took longer to complete the writing of articles.

In contrast, although interviewees from schools with both high and low proportions of children receiving FSM reported challenges related to the resources needed for the programme, this issue was emphasised more in schools with lower FSM rates, although the reason for this was not clear. An additional challenge in a school with a low rate of FSM stemmed from the fact that it was part of a small and isolated community. Their mentor felt that this meant pupils sometimes struggled to understand events taking place outside their tight-knit community. As a result, pupils had a narrower range of topics on which to base their article writing.

> *'It's a smaller school. They are closed off from a lot of things like outside in the world... So it's been a bit more difficult, like trying to get children invested in the news for instance. Like they know kind of like the local area… but initially they didn't have that awareness of outside.'* (YJA mentor 04)

Teacher's experience of journalism

Teachers with a prior interest or experience in teaching journalism were highlighted by YJA mentors and members of the SLT as being particularly good at supporting children to incorporate learning and skills from YJA into the curriculum. These teachers also articulated their own enthusiasm for taking part in the programme. One teacher described the benefit that learning journalistic techniques and styles from the mentor had on their teaching:

> *'It's been really good to learn things from him as well, sort of techniques and styles of journalistic, you know, writing... Developing skills for when I then teach and you know phrases that the children are now familiar with to use, and things to look for and features to look at more closely when doing that work.'* (Teacher 03)

A mentor reported that teachers who had more limited experience of teaching writing, or did not see it as their strong point, were less immediately engaged in terms of delivering YJA activities in the classroom and incorporating learning into lessons.

> *'I work with another teacher... she's not present. So she's not open to the idea of what we are doing and how we are doing it. I think it's a shame because it's a good skill and tool to have and*

use, even if it's just one thing you teach them, they get up and use it themselves. Then I have others, one is a trained journalist... and she is all over this.' (YJA mentor 06)

Usual practice

Survey data indicated that teaching journalistic skills through article writing was an existing part of the curriculum for most schools in the evaluation (see Table 11). In the past year, similar proportions of control (51%) and intervention schools (62%) had taught writing newspaper articles outside of YJA time. Case study interviews with class teachers and the SLT from these schools also confirmed that reading newspapers and article writing were part of their existing writing curriculum.

Table 11: Survey responses regarding journalism-related classroom practices

Intervention: Which of the following activities have you done with your class, outside the eight Young Journalist Academy teaching days?		
Control: Please tick which of the following activities your class has taken part in this year.		
	Intervention n (%)	Control n (%)
Reading newspaper articles	20 (57%)	16 (49%)
Writing newspaper articles	22 (63%)	17 (51%)
Listening to radio programmes	8 (23%)	5 (15%)
Watching the news/news clips	21 (60%)	17 (51%)
Total respondents	35	33

However, teachers and members of the SLT in intervention schools noted that a key difference between the journalism-related tasks the pupils did as usual practice and YJA-related tasks was that pupils would not usually be given the chance to do activities such as radio and TV/film production. As a result, they felt that YJA had 'brought journalism to life' and engaged pupils with the real world, describing it as a positive addition to the existing curriculum.

> *'I know in Year 6 we do newspapers and things, but... it's not the same. We look at a newspaper and analyse it and it's not bringing it to real-life... There's that clear link between what they do in the classroom with the project and how that can affect them in the future... I think that's the difference, it really brings it to life.'* (SLT member 03)

Implementation

This subsection discusses factors affecting the fidelity (RQ1) of implementation of the YJA programme.

Fidelity to the delivery model

In order to deliver the YJA programme as intended, classes needed to receive eight YJA teaching days, during which a YJA mentor facilitated teaching of the programme, as well as YJA skills being developed in the class outside these teaching days. In terms of the eight teaching days, the aim was for YJA mentors to spend a reasonably balanced amount of time with each class, working on each different activity (article writing, radio production and TV/film production).

Among intervention schools completing the survey, the most frequently selected number of programme days spent on article writing was four (46% of respondents, 16 of 35 surveyed), while two days was most frequently reported for TV (60% of respondents, 21 of 35 surveyed) and radio production (77% of respondents, 27 of 35 surveyed). However, in some schools most time was spent on one activity, with 29% of survey respondents reporting that their class had spent five or more of the eight days solely on article writing, and a small proportion (six percent, two of 35 surveyed) reporting that they spent no time on TV/film production (the reason for this was not apparent). This data is summarised in Table 12.

Table 12: Number of teaching days (out of eight) spent on each YJA activity, as reported by survey respondents

Activity	Number of YJA days spent on activity	*n* (percentage of respondents)
Article writing	2 days 3 days 4 days 5 days 6 days 7 days	4 (11%) 5 (14%) 16 (46%) 6 (17%) 3 (9%) 1 (3%)
Radio production	1 day 2 days 3 days	4 (11%) 27 (77%) 4 (11%)
TV/film production	0 days 1 days 2 days 3 days	2 (6%) 11 (31%) 21 (60%) 1 (3%)

Notes. *n* = 35; Due to rounding, some figures may add to slightly above or below 100%.

It was also intended that a 'core team' of pupils was set up to edit and proof articles written by the class, as part of a peer editing process, before final articles were sent to the teacher to review and upload to the YJA website. In schools where a core team was set up, there were three approaches to selecting this team:

(i) selected by teachers based on attainment levels;

(ii) selected by teachers randomly;

(iii) selected through popular vote by class pupils.

Core teams were not selected in all schools, however, and a mentor said that articles submitted without input from the core team required more edits. However, even when schools did select a core team it was not always possible for them to fulfil the role intended by the programme, because the team did not always collaborate effectively or have the skills to act as editors.

In addition, time constraints meant that the process of editing had to be severely curtailed or was absent. For example, in a school with a core team, pupils struggled to both complete their own work and proofread their peers' articles, and the mentor described how they then had to take over the proofing of articles, to allow the pupils to finish their

own work. In another case study school, the teacher felt that pupils selecting their own core teams was good for generating ownership of the role, but also led to the core teams being less diverse in terms of attainment. The teacher felt that having core teams with less diversity resulted in some members of the core team struggling with confidence in their roles, and that more diverse core teams would help to ensure that pupils with high attainment in literacy (perceived as more confident) were not all in one group, and hence allow pupils with lower attainment in literacy to be supported by their more confident peers.

> '[For] everything we've done there has been different groupings, and I think that has caused a few problems. Just in terms of some children not working as hard as others, or not feeling as confident within that group. So I think if we had a mixture of ability levels and confidence levels... then that might have helped the teams to be more successful.' (Teacher 01)

Implementation barriers and facilitators

Following analysis, the logic model was updated to reflect the IPE findings (Figure 8).

Figure 8: Revised logic model for the Young Journalist Academy (YJA)
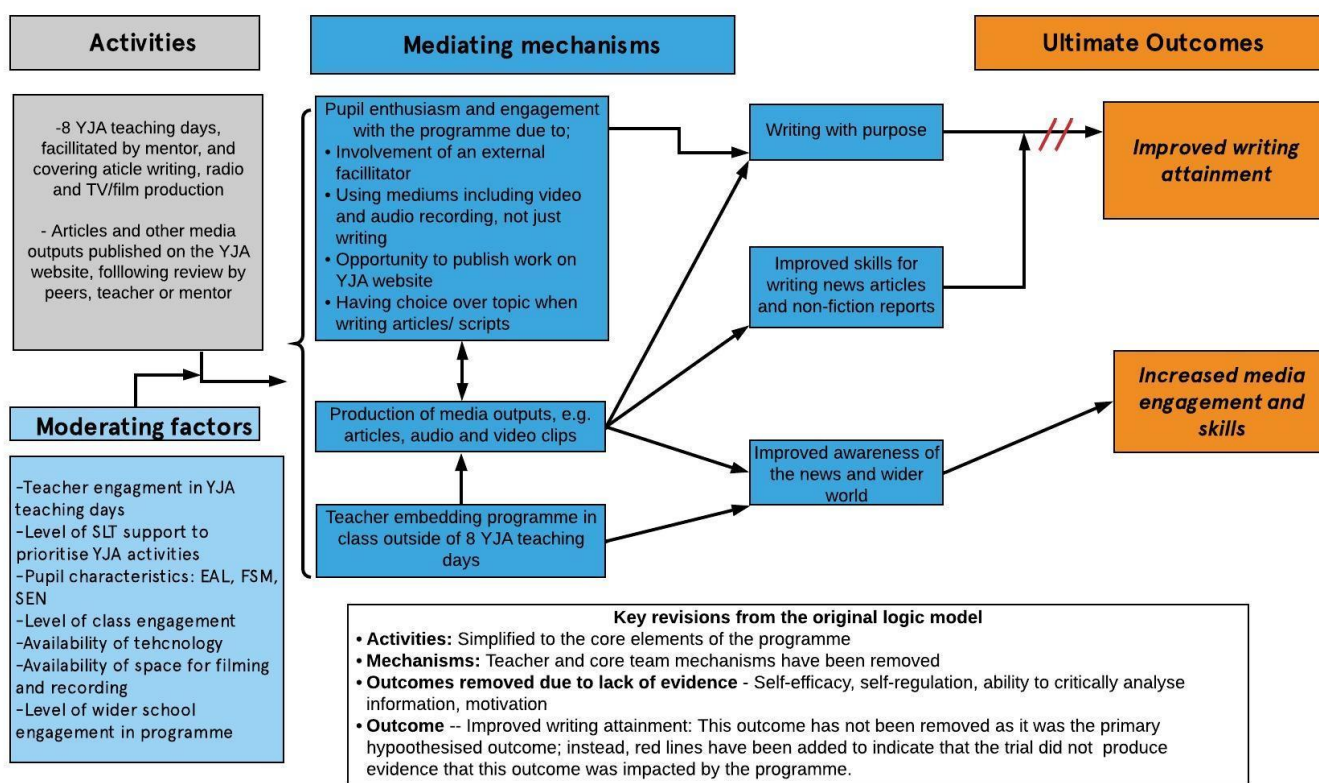
This subsection discusses five overarching factors that were identified as acting as barriers and facilitators to the implementation of YJA: SLT support; teacher engagement in the sessions; teachers prioritising YJA outside the sessions; pupil engagement; and resource availability.

SLT support

Engagement and support from school SLTs affected the delivery of YJA in two main ways:

1. by allowing teachers to dedicate lesson time to YJA activities outside the eight teaching days;

2. by showcasing YJA in the wider school, and encouraging greater staff and student engagement.

The majority of survey respondents (66%) stated that the SLT was moderately to highly influential in their decision to do journalism-related activities outside the eight allocated teaching days. This was reflected in case study data, which indicated that SLT support to use standard literacy time for article-writing, as well as integrating YJA skills into other areas of teaching, was important for ensuring that YJA learning went beyond the eight teaching days. A teacher reported that their headteacher had not allowed them to use additional lesson time for YJA activities, which had limited their ability to embed YJA learning outside the days when the YJA mentor visited the school.

> *'...we don't always have the time to do, or to do as well as we could like, the interim tasks. Because obviously we have to cover a lot of different fiction, non-fiction genres of writing. To fit everything in and do that as well, the head didn't really want us to use our literacy time to do the interim tasks. So fitting that in was a bit of a squeeze.'* (Teacher 04)

Teacher engagement in the sessions

As outlined in the logic model in Figure 8, one moderating factor influencing implementation was teacher engagement during the eight days that the YJA mentors were in school. Observation notes showed that teachers who took an active role in YJA sessions played an important role in providing guidance to pupils about their work, and supporting pupils to stay on task, instead of using the YJA sessions as an opportunity to step out of class or complete administrative tasks. More actively involved teachers were also more able to reinforce learning from YJA once the sessions with mentors were completed. Interviews with mentors highlighted the difference between an engaged teacher and a disengaged teacher in the programme, suggesting the teachers who were engaged in the programme had a better sense of the work pupils were doing, and were better placed to continue developing pupils' YJA skills in the future.

> *'[The schools] where you've got a teacher that's engaged in the project are the ones with the kids [that] are more engaged in the project... So, the teachers who sit scribbling notes that they are taking in what you are speaking, and the ones who really delve into it themselves and ones working intensely with groups... They are the ones where you see... more content produced and you see good quality articles getting written.'* (YJA mentor 04)

> *'Whereas there are other teachers in other schools, that when I get to the school they show me to the classroom and they go, "and I'm going to be in my office all day", and they don't join in, and they don't listen. So you know that when you leave they are not going to know what we were doing and they are not going to carry on the legacy.'* (YJA mentor 06)

The relationship between teachers and the class mentor was another factor affecting implementation. Data from the survey showed that 77% of teachers (27 of 34 surveyed) found the mentor easy to work with (with 54% strongly agreeing and 23% agreeing). Findings from the qualitative interviews were in line with the survey data in that both mentors and teachers generally described their working relationships as positive. A teacher described how they enjoyed working with their mentor due to their enthusiasm and ability to engage pupils.

> *'Brilliant, yeah, she's really full of energy from minute one in coming in. She is really competent in what she does and she makes it fun. She's built up a relationship with the children really quickly. Yeah, she's been really great.'* (Teacher 05)

Teachers embedding YJA outside of the sessions

In the survey, 64% of respondents reported completing YJA activities with their class more than once a month, outside of the scheduled teaching days, whilst 26% of respondents did so less than once a month (3% of respondents reported that it varied). The most commonly completed activities included writing newspaper articles (63% of survey respondents), watching the news or news clips (60%) and reading newspaper articles (57%). In addition, teachers and pupils reported that children completed work related to YJA at home or at lunchtime, including researching and writing articles and listening to podcasts. Teachers felt that making time for YJA activities helped reinforce learning from YJA sessions, enabling pupils to build on the skills that they had learnt and to feel that YJA was something that was integrated into standard learning, rather than a one-off event.

Through the case study data, factors that appeared to help teachers to prioritise the completion of YJA activities outside the scheduled teaching days were SLT support (as described above), teachers taking the time to watch the mentor's delivery so that they could model their own lessons on them, and teachers 'buying-in' to the programme themselves. However, some teachers reported real difficulties with finding the time to do so, with teachers even reporting that just taking part in the scheduled YJA days meant that they had to reduce the amount of time spent on creative subjects, such as art and music.

> *'So today, I've missed obviously quite a few lessons that the other Year 5 classes are teaching, and then I have to try to fit those into my timetable...it's not like this project has replace[d] my English lessons or anything... we've ended up having to miss some of the nicer subjects... subjects like music and things and art have kind of been left out.'* (Teacher 01)

In order to disseminate learning from the programme more widely, some of the case study classes delivered assemblies about YJA, to showcase their journalistic outputs and disseminate the skills they had learnt to the rest of the school, or set up newspaper article suggestion boxes to allow other classes to feed into the YJA work. Teachers also reported passing on YJA learnings to other Year 5 classes within their school, to ensure consistency and shared knowledge. However, there were also schools that struggled to disseminate YJA beyond the class that took part, where teachers said they had limited time to spend on YJA between scheduled sessions.

> *'We did try and encourage any ideas for articles. That wasn't very successful and again that's probably down to the fact that me being two days a week, I haven't always been able to follow that up to encourage that more.'* (Teacher 02)

Pupil engagement in the sessions

Overall, 69% of survey respondents (24 of 35 surveyed) said that their pupils found YJA engaging, with 46% 'strongly agreeing'. This was echoed by teacher and pupil comments about enjoying YJA lessons because it provided the opportunity to choose what to write about, learn new skills, and publish work online. However, in this *Implementation* section, the focus is on how aspects of implementation affected levels of engagement; more detail about the way that the intervention fostered engagement can be found in the *Mechanisms and Impact* section below.

Observation notes recorded that during YJA sessions the majority of pupils were on task, engaged and absorbed, but that in some classes, the engagement of some pupils could dwindle and result in distracting or off-topic discussions, particularly as the day went on. Where this was the case, the reasons observed for this included: use of long PowerPoint presentations; groups of children struggling to reach agreement about what topic to base their work on; and children having to wait to record or film, or to get guidance from their teacher or mentor. In particular, as reported in teacher interviews and observation notes, children could struggle to maintain concentration when they were asked to write articles or scripts for extended periods of the day.

> *'A lot of it to begin with was all about writing, and a full day of writing is tough. So I think perhaps at the beginning there still needs to be some kind of practical more teamwork aspect of things, and the chances for them to get up and move around and do different things that aren't just working in your group and writing.'* (Teacher 06)

This was also reflected in a couple of comments by pupils that they could find it difficult to work when the class was distracted and that it might be better to deliver YJA in half days, so that the teaching day did not feel as long.

To minimise disruption and support pupils to engage, teachers suggested that spending more time doing practical and interactive tasks such as radio and TV/film production, as well as taking regular breaks, could be helpful. Observation notes also identified that breaking delivery into a series of short sessions, each with a specific focus, appeared to maintain engagement throughout the day. In addition, one teacher explained how they had assigned pupil 'talk partners' to some pupils who needed to talk as they worked. This enabled these pupils to have an assigned person to chat to, without disrupting the wider class.

Resource availability

While 71% of survey respondents (25 of 35 surveyed) agreed that the required technology was available to deliver the YJA programme, 20% disagreed (9% neither agreed nor disagreed). Observation notes made in some schools, as well as comments by interviewed teachers, indicated that limited availability of laptops or tablets (i.e. less than one per two pupils) and other equipment meant that there were occasions where children had to wait around, either while other members of their group used the laptop, or to record with the mentor. At times, this could result in pupils losing interest and engaging in distracting behaviour. There was some confusion over who was responsible for providing the equipment needed for the YJA programme. A number of teachers reported being surprised by the lack of equipment provided by mentors as part of the programme.

> *'I was a bit surprised that there was a limited amount of equipment... they're having to wait quite a long time to use the equipment to go and record their interviews and their videos... Even if there were four microphones instead of just the one. I think the process of all of these different groups getting to record it would just make the whole thing run quicker.'* (Teacher 05)

On the other hand, some mentors were surprised by the limited equipment schools had, and the fact that they often had to rely on pupils using their own equipment (e.g. iPads) at home. Mentors felt schools should be equipped with the necessary equipment prior to the programme starting. Some schools already had access to resources such as iPads (used for recording) and laptops for writing articles; however, not all laptops provided by those schools had the recording or editing capabilities required to complete some YJA activities.

Teachers reported that it was expensive for schools to take part in YJA when equipment needed to be purchased. Ideally, they felt that the programme should provide access to a suite of laptops with Word installed, and tablets/cameras for recording. They felt that one laptop per pupil would speed up the process, particularly as it would allow articles to be uploaded more efficiently. It was also reported that at one school, they only had laptops from day six, and prior to this it was difficult to get articles written.

Classroom space was not generally a barrier to undertaking the YJA activities, as 85% of surveyed teachers (30 of 35 surveyed) thought the required space was available to deliver the programme effectively. Nevertheless, one teacher reported finding it difficult to find a quiet space to record. In this case, the mentor would usually take small groups of pupils out to record, while the teacher would oversee the article/script writing in the main classroom. There were also examples of this in the observation data; for instance, in one case, there was a group who could not find anywhere in the school to record, and ended up having to use a small store cupboard.

Mechanisms and perceived impact

This section discusses evidence about the intervention's mediating mechanisms and outcomes, after which a revised intervention logic model is presented. Three mediating mechanisms were identified, but not tested, through interviews and observations:

1.   pupil engagement with the programme;

2.   production of media outputs;

3.   the teacher embedding the programme in their class outside the eight YJA teaching days.

These are hypothesised to lead to one intermediate outcome focused on media and journalism:

- improved awareness of the news and wider world

and two intermediate outcomes focused on writing:

- writing with purpose

- improved writing skills.

If this pathway is successful, it has the potential to result in two ultimate outcomes:

1. increased media engagement and skills

2. improved writing attainment.

A number of broader reported outcomes were also identified, such as improved verbal communication and social skills, which are detailed towards the end of this section, but are not included in the revised logic model (Figure 8) because they were not seen as part of a pathway to the ultimate outcomes.

Mediating mechanisms

Pupil engagement with the programme

As indicated in the *Implementation* section, the majority of surveyed teachers felt that, overall, their pupils found YJA engaging, though some pupils in some of the case study schools struggled to maintain attention for the whole day. Though there were pupils who said that they preferred writing stories (including both boys and girls), pupils were generally enthusiastic, describing YJA as 'fun' and more enjoyable than English, with one child stating:

> *'I kinda wish it was everyday we did this.'* (School 01)

Teachers also reported that pupils enjoyed the programme. A teacher described how enthusiastic and engaged the pupils were about YJA and how much they looked forward to the sessions.

> *'Yeah, they have been really successful. The children are always engaged as soon as [the YJA mentor] walks in. Their faces light up and they cling onto every word. Sometimes they get a little bit over excited in some of the sessions, but I think that's just their enthusiasm.'* (Teacher 03)

There were two elements of the programme identified that facilitated pupil engagement. The first element was the fact that the programme was delivered by an external provider. Pupils commented that they liked their mentor because they found them encouraging, good at explaining things and knowledgeable about journalism. Staff interviewees also reported that mentors who were positive, enthusiastic and had experience working with children were good at motivating pupils to want to learn. In some cases, this resulted in pupils voluntarily taking YJA work home to complete. Another benefit of having an external mentor was that pupils were described as valuing praise from them because the mentor was separate to the school. It was felt that because the mentor did not have prior knowledge of struggles or difficulties that pupils had, the pupil could develop a fresh relationship with the mentor, based on the work they were producing as part of YJA.

> *'I think just a new person, because... it might be that their teacher knows they struggle with writing and he [the mentor] doesn't know any of that. So they've just been allowed to get along, and he's really positive. Like every piece of work they bring up to him he's like it's the best work he's ever seen. That's how he treats all the pieces of work.' (Teacher 03)*

The second element was that the program involved a range of different activities. One of the key reasons that pupils gave for enjoying YJA was that they could 'learn new things', including how to use technology, conduct research and produce articles, podcasts and TV/film clips. Interviewees noted that the range of different activities involved in YJA gave pupils who might struggle with more traditional teaching approaches a way of engaging. In one example, a mentor described how a pupil with SEN, who members of staff felt would not engage with YJA, ended up recording a radio clip, which was published on the YJA website.

> *'There's a child in this class... he's incredibly low ability, and when I met him, the TA [teaching assistant] said to me, he will not join in, this will be totally beyond him… at the end of the first day*

he recorded his first radio show. He interviewed members of staff... He had written a script. He had written questions... For him to walk away from the classroom, from "he won't do anything" to being broadcast on the internet... is phenomenal.' (YJA mentor 05)

Production of media outputs

Across data sources, there was clear evidence of children working to produce a range of media outputs. Survey responses indicated that all classes had engaged in article writing and radio production and 91% in TV/film production. Children also reported writing articles and scripts on a range of topics, including popular culture (e.g. film and book releases), news and politics (e.g. plastic pollution and the Notre Dame fire), sport (e.g. football and Formula 1) and local events (e.g. Christmas light switch-ons). Observation notes also showed that lessons facilitated by YJA mentors involved the writing of articles and scripts for future recording and filming, as well as the actual recording and filming of radio and video/television clips. There was also evidence from across the observations of mentors providing guidance in order to refine media outputs, with mentors providing tips such as asking questions to the audience in order to keep them engaged, and avoiding filming with light behind them as this created shadows. Where the outputs were of high quality, there was the opportunity to publish them on the YJA website. The administrative data showed that intervention schools published, on average, 23 media items on the YJA website (range 7–46), demonstrating that children in all schools were not only producing media outputs, but media outputs of a publishable standard.

Teachers embedding YJA learning

The original logic model had two teacher-focused mechanisms through which the intervention was anticipated to have an effect. These were:

1.   fostering collaboration amongst teachers;

2.   providing more speaking and listening opportunities.

The evidence gathered from the IPE indicated that the key mediating mechanism for teachers was in fact whether YJA learning was happening outside of the eight YJA teaching days. Information on the ways that teachers embedded YJA outside the teaching days, as well as some of the facilitators and barriers to doing so, is included in the Implementation section (especially the subsection Teachers prioritising YJA outside the sessions). The potential importance of this as a mediating mechanism was reinforced through case study interviews with mentors, in which they reported that they observed a drop in impact of YJA when momentum was not maintained between their visits. Ensuring teachers have sufficient time to incorporate YJA into existing literacy lessons and activities may enable the intervention to continue to be impactful on children's learning when the mentor completes the sessions.

Media-focused intermediate and ultimate perceived outcomes

Improved awareness of the news and wider world

The intermediate outcome identified related to the media was that pupils showed increased interest in the news and wider world. Survey data showed that 74% of teachers thought that YJA had a positive impact on pupils' engagement with culture and the wider world (with 37% of teachers saying this impact was 'very positive'). From the case study data, there was also evidence that YJA had helped engage pupils in current affairs and the news, and increased pupils' awareness of what is happening in the world outside their school and local area.

> *'In terms of the biggest thing actually for my children here and a little… village is the world news out there and other things they might be interested in. I think that's the* biggest *thing for them is actually finding a little bit more on what's going on away from their doorstep.' (SLT member 04)*

Increased media engagement and skills

Building on the improved awareness of the news, there was some evidence that longer lasting changes were taking place for pupils in relation to media engagement and skill development. For instance, at one school, staff members reported that there had been increased uptake of the school's after school Journalism Club from pupils who had taken part in YJA. There were also instances reported of the programme increasing pupils' interest in pursuing journalism as a potential career option, which they had not considered prior to their exposure to YJA.

> *'You know, you don't just have to be good at writing, so I can see for those children that can be really empowering in terms of actually you know, I could have you know a career in this field. Just because I'm not the best writer in my class doesn't mean I couldn't go and work for a newspaper for example.' (SLT member 01)*

Secondly, in terms of skills, teachers noted that, through YJA, pupils had learnt to better differentiate between facts and opinions, had developed skills for interviewing, and had improved their verbal communication when reporting factual information, for instance, as part of radio shows. Observation data and discussion with pupils also indicated that children had developed skills for researching factual information online and understood the importance of ensuring that the information they reported was accurate. Finally, interviewees reported that there were improvements to children's writing of news articles, which is covered in the following subsections.

Writing-focused intermediate and ultimate perceived outcomes

Writing with a sense of purpose

Case study data indicated that YJA had helped pupils, particularly those described as reluctant writers, to find purpose in their written work. One element that facilitated this was that pupils' articles were published online. Discussions with pupils and case study interviewees indicated that this motivated pupils, and gave them a sense of pride that their work was published online and could be read by other people.

> *'For me it was also an exciting project because it gives purpose to their writing. You know, they sit and they do all of this writing in their exercise books, but it's more motivating for them to know that it would be read by somebody from outside the school or it might go online. That I think is a really important factor.' (SLT member 06)*

Another aspect of this was that the sessions contained a combination of practical and written activities. This was described by a teacher as important for giving children who might struggle with typical teaching approaches a reason for writing: without producing a written script, they could not go on to produce the radio or TV show.

> *'It has engaged some of the lower children and because it has been hands-on and working with materials, it's helped them to see that writing does have a purpose. If I don't write this script or these interviewer questions, then I can't go on and produce a radio or TV programme. So, there has to be some writing and there has to be some planning.' (Teacher 05)*

In addition, the programme gave pupils the freedom to choose what to write about. Pupils described this as one of their reasons for liking the YJA programme, in contrast to other English lessons in which the topic was set for them by their teacher. Interviewees also identified this freedom as a key reason that pupils, particularly those who they might think of as reluctant writers, had found purpose in writing. This was described as being particularly the case for boys, who had started to understand through YJA that writing was more than just stories, and was an important part of communicating information.

> *'When you give them something that they're interested in and they can run with and they're allowed to produce what they want. So I think it's opened their eyes and thinking, "I'm actually enjoying writing because I'm interested in this".' (SLT member 03)*

> *'I think particularly the boys as I said have developed more of a focus for writing and an understanding that writing isn't always about the fairy stories and the stories, and that they have perhaps been developing a pride in their writing.' (SLT member 02)*

Improved writing skills

The other intermediate outcome related to writing was that pupils demonstrated improved writing skills. Examples of this included using specific skills that form part of the YJA programme, such as the 'six tools' (i.e. who, why, where, what, when and how), and the 'three Cs' (i.e. clear, concise and correct) within their writing. These skills were reinforced at several points during the observed lessons, and were identified by pupils during discussions as one of their key learnings through YJA, and they were keen to demonstrate that they knew what each of the 'six tools' or 'three Cs' were.

Teachers felt these techniques had supported their classes to develop skills for writing newspaper articles and re-ports, both within YJA and as part of literacy lessons.

> 'It would be like you know, the six tools and sort of looking at my own practice and watching them and thinking, we need examples of those and we need to really break down, who, why, what, where, when. Then provide the children with lots of opportunities to develop that themselves. So, small steps learning in terms of writing an article.' (Teacher 02)

Case study data also showed that teachers felt pupils had developed broader writing skills, such as writing for a spe-cific audience, improved use of punctuation, tense changes and reported speech, as well as becoming more used to using thesauruses and dictionaries to guide their writing. A class teacher described how pupils' vocabulary, terminol-ogy and writing in English lessons had improved.

> 'So it's just been really, really nice to see and I think the vocabulary that they're learning and the news, and the correct terminology has really improved. It's also helping with their writing in English lessons.' (Teacher 01)

Improved writing attainment

The ultimate outcome in the logic model in terms of writing, is improved writing attainment. Of the teachers surveyed, 69% thought that YJA had a positive impact on pupil's writing (with 26% thinking it had a very positive impact). Inter-viewees echoed this, with one class teacher describing how the YJA program had led to an improvement in pupil's writing, particularly their factual writing:

> 'The writing has come on. I think typically with news reports and non-fiction writing, that's where I've seen the biggest improvement. They struggle creatively when it comes to writing and they struggle with things like that. Obviously, all of the news reporting that they are doing is all factual, so that has helped them with their writing.' (Teacher 03)

There was, however, also scepticism about the perceived benefit of the intervention on pupil's writing. As well as 31% of survey respondents reporting that they did not believe that YJA had much of an impact on pupil's writing, there were case study interviewees who reported being uncertain about whether YJA would have improved writing attain-ment, because they felt that the programme did not focus on formal writing or spelling and grammar. Instead, these teachers highlighted that the programme may have had more of an impact in terms of engaging more reluctant writers and increasing writing confidence.

> 'I would be very surprised if it's had a massive effect on their actual outcomes. I won't be surprised if some of the more reluctant writers are now more willing to write. I think the engagement would have caused that, but I will have to wait and see until a little later if it has.' (Teacher 05)

Other perceived outcomes for pupils

Two other broader outcomes were also identified: improved verbal communication skills and confidence, and en-hanced social skills. Because there was nothing in the data to clearly link these to one of the ultimate outcome path-ways, they are detailed here, but not included in the logic model.

The survey found that 77% of teachers (27 of 35 surveyed) thought that YJA had a positive impact on pupils' confi-dence, and 80% felt the programme improved communication skills (28 of 35 surveyed). Where interviewees spoke about increases in confidence and communication skills, this was often linked to verbal communication. Interviewees reported that taking part in YJA had improved both pupils' confidence in speaking out loud as well as their clarity of speech. Teaching staff described how pupils, particularly those considered shy or less confident, and who would not usually put themselves forward for public speaking, had recorded radio shows as part of their group, spoken in front of their class, and contributed to assemblies in which the YJA programme was introduced to the wider school. In one example, a pupil with a speech impediment, who was described as initially hesitant to speak, now loves joining in les-sons and sharing their work in front of the class.

> 'My class did an assembly to explain... what the project is all about and what kind of things they've been doing... There were children who wouldn't usually offer to speak. I mean it was in front of

quite a lot of people... and they were just confident and clear and it was amazing to see. (Teacher 04)

'The radio show has helped them with the communication skills and to speak aloud... They have to practise it and then they have to record it, and then they would listen back and see where they have not been very clear... that process has really got them thinking about how to speak in front of an audience.' (Teacher 03)

In terms of social skills, 74% of survey respondents felt that YJA had a somewhat or very positive impact for their pupils. There were also some comments in interviews about how working in a team had helped develop pupils' social skills. For instance, a member of SLT described how the Year 5 class involved in YJA had initially struggled working as a team, but over time, pupils had become more familiar with working and collaborating as part of a group.

'Over the project [the class teacher] has seen an improvement in the children in terms of their teamwork. I know initially in the first session it would be like, right get into groups and 'I want to be with my friend', and 'I won't work with that person'. Whereas now, you... say right, get into your groups and they will just go off and do that, so that is definitely an improvement.' (SLT member 01)

Cost

Delivery of the YJA intervention cost approximately £944 per school for the year it was delivered, and the majority of costs were realised in the first year. Expenses reported included photocopying and provision of lunch to the YJA mentor. Programme fees were paid to the developer. As noted in the IPE findings, some IPE case study schools indicated that it was necessary to purchase expensive equipment such as laptops in order to implement the programme. However, among the schools sampled to take part in the cost evaluation, none reported needing to purchase equipment in order to participate.

To calculate the total cost per pupil over three years, we assumed the number of pupils would cumulatively increase from 25 pupils in Year 1, 50 in Year 2 and 75 in Year 3. We assumed 25 pupils as this was the number of treated pupils. Based on these assumptions, the total cost per pupil per year over three years is £13 (set out in Table 13). The cumulative cost breakdown is set out in Table 14 (all figures are rounded to the nearest pound).

Table 13: Cost of delivering Young Journalist Academy (YJA)

Item	Type of cost	Average cost (£) (minimum, maximum)	Total cost over 3 years (£)	Total cost per pupil per year over 3 years
Expenses	First year	6 (0, 22)	18	0
Programme fees	First year	1,250 (1,000, 1.500)	1,250	17
Total		1,256	1,268	17

Source: Interviews with teachers from cost evaluation case study schools (*n* = 4).

Table 14: Cumulative costs of YJA (assuming delivery over three years)

	Year 1	Year 2	Year 3
Cumulative costs	£1,256	£1,262	£1,268

Source: Interviews with teachers from cost evaluation case study schools (*n* = 4).

Delivery

In order to deliver the YJA programme as intended, teachers and teaching assistants were required to deliver eight YJA teaching days. On average schools delivered 45 hours (or 4.5 days) per school per year. The reported hours of

intervention delivery ranged from 0–66 hours. In addition, respondents reported spending, on average, four hours outside of the eight programme days delivering YJA activities per school.

Resources

As reported in the IPE findings, the majority (71%) of survey respondents indicated they had the necessary technological equipment to implement the programme. In line with this finding, the cost case study schools all reported using technology they already possessed to implement the programme (e.g. laptops, tablets and interactive whiteboards) and did not need to make any further technological investments to participate in YJA. As such, the cost figures above do not reflect any purchases of technological resources. As reported in the IPE findings, some schools did not have sufficient technological resources to implement the programme as designed and either had to make purchases or found ways to deliver without.

Preparation

There was some administrative time associated with supporting the delivery of the intervention. Cost survey respondents reported spending on average one hour preparing to deliver YJA activities outside of the eight programme days.

Conclusions

Table 15: Key conclusions

Key conclusions
1. Children in schools that participated in YJA made the equivalent of two months' less progress in writing as measured by the WAM, on average, compared to children in other schools. This is our best estimate of impact, which has a moderate to high security rating. However, as with any study, there is uncertainty around the result: the possible impact of this programme ranges from four months' less progress to positive effects of one additional month of progress.
2. Among pupils eligible for FSM, those in schools that participated in YJA made the equivalent of three months' less progress in writing, on average, compared to children in other schools. These results have lower security than the overall findings because of the smaller number of pupils in this group.
3. There is no evidence that YJA had an impact on writing self-efficacy or writing creativity (ideation) as measured by the WSEM.
4. Findings from the IPE indicated that teachers perceived the programme to have a positive impact on pupils. Of the teachers surveyed, 69% thought that YJA had a positive impact on pupils' writing. However, some teachers were uncertain about whether YJA improved writing attainment, though these teachers said the programme may have had more of an impact on engaging more reluctant writers and increasing writing confidence.
5. Among teacher survey respondents, 74% thought that YJA had a positive impact on pupils' engagement with culture and the wider world, and that there was some evidence that longer lasting changes were taking place for pupils in relation to media engagement and skill development. Some teachers also felt that YJA had a positive impact on pupils' confidence, and that the programme improved communication skills. That said, some teachers found it difficult to reconcile the amount of time required for YJA with teaching the school curriculum and found it challenging to further embed the programme outside of the sessions.

Impact evaluation and IPE integration

The goal of this impact evaluation was to answer three key questions:

1. What is the effect of participating in the YJA over the course of one school year on pupils' writing skills?

2. What is the effect of participating in the YJA over the course of one school year on pupils' writing self-efficacy?

3. Does participating in the YJA over the course of one school year have an impact on pupils' perception of their own capacity to generate ideas?

The IPE focused on addressing research questions related to the implementation, delivery and perceived impact of the intervention. The IPE was also used to identify how the YJA compares to usual practice, and to determine the cost to schools of delivering the intervention.

Evidence to support the logic model

The original logic model for this intervention was revised to reflect findings from the evaluation. In the revised logic model, programme activities are reduced to the core activities of the eight YJA teaching days and publication of media outputs. As some case study schools did not select a core team and some found it difficult for the core team to fulfil their responsibilities, this activity was omitted from the logic model.

The programme mediating mechanisms were regrouped into three broader categories:

1. pupil engagement with the programme;

2. pupil production of media outputs;

3. the teacher embedding the programme in their class outside the eight YJA programme days.

Based on findings from the IPE, new intermediate outcomes were identified: one focused on media and journalism:

- improved awareness of the news and wider world;

and two focused on writing:

- writing with purpose;

- improved writing skills.

If this pathway were successful, it could result in two ultimate outcomes:

1. increased media engagement and skills

2. improved writing attainment.

The impact evaluation results for this trial found no significant effect of the YJA intervention on pupils' writing skills, the primary outcome. The overall effect on the WAM (the primary outcome measure) was small and slightly negative. It was not statistically significant. The impact evaluation results on the secondary outcome measures, the WSEM and the ideation score, were small, positive and also not statistically significant. Overall, there was no significant impact of the YJA on any outcomes identified in the logic model. In accordance with these findings, the logic model was revised to remove self-efficacy as an outcome. As improved writing attainment is a primary aim of the programme, this outcome was retained in the model but with an indicator that there was no evidence to support this outcome.

The IPE data showed that pupils in case study schools engaged with the programme and produced media outputs. This was reinforced by programme administrative data which showed that intervention schools published, on average, 23 media items on the YJA website (range 7–46), demonstrating that children not only produced media outputs, but these outputs were of a publishable standard. Teachers reported that they perceived the intervention had a positive impact on the writing of their pupils – particularly in regard to sense of purpose and writing skills. The evaluation did not explicitly evaluate the programme's impact on the other outcomes identified in the original logic model: motivation, self-regulation and ability to critically analyse information. However, there was also little evidence from teacher interviews that there was a perceived impact on these outcomes. All three were omitted from the revised model.

Two other broader outcomes were identified through the IPE: improved verbal communication and social skills. However, there was nothing in the data to clearly link these to one of the ultimate outcome pathways, so they were not included in the revised logic model.

The original logic model also had two teacher-focused mechanisms through which the intervention was anticipated to have an effect. These were (1) fostering collaboration amongst teachers; and (2) providing more speaking and listening opportunities. The evidence gathered from the IPE indicated that the key mediating mechanism for teachers was in fact whether YJA learning was happening outside of the eight YJA teaching days. Slightly more than half of survey respondents (64%) indicated they were engaging in YJA-related activities at least once a month outside of the scheduled teaching days, and 26% of respondents indicated they were doing so less than once a month. Three factors were identified as potential facilitators or barriers for embedding: SLT support, teacher buy-in to the programme, and teacher engagement in the YJA sessions. Teacher outcomes were not collected as part of the impact evaluation.

Interpretation

YJA is a novel intervention that aims to create pupil interest in journalism and improve writing skills through the experience of setting up a 'newsroom' in primary schools. The programme draws upon several interventions that have some evidence of improving writing in pupils (Years 4–12) (Graham & Perin, 2007), but there is little available evidence as to the causal impact of this type of journalism programme. As such, this evaluation set out to understand YJA's impact on writing attainment (the primary outcome), as well as writing self-efficacy and ideation.

> *The impact evaluation did not find a statistically significant effect of the YJA on either primary or secondary outcomes. The measure of writing indicated that the average outcome for the intervention group was marginally worse than for the control group (but not statistically significant). For writing self-efficacy and social skills, the average outcome for the intervention group was marginally better than for the control group, but again, the difference is not statistically significant.*

However, some data from the IPE indicated that teachers perceived the programme had positive effects on pupils. Of teacher survey respondents, 69% thought that YJA had a positive impact on pupils' writing and some interviewed

teachers perceived improvements in writing-related intermediate outcomes: writing with a sense of purpose and writing skills. Conversely, 31% of survey respondents felt YJA had a limited impact on writing attainment. Some case study interviewees indicated that one limitation of the programme was the lack of explicit focus on formal writing and spelling and grammar. The interviewees also felt that the programme may have had a bigger impact on engaging reluctant writers and improving pupil confidence. Teacher survey respondents also observed positive impacts on pupils' engagement with culture and the wider world and communication skills, but these outcomes were not part of the impact evaluation.

Overall compliance with the programme was very high. All but one school met the minimum compliance threshold, which required that schools deliver all eight sessions and upload a minimum of ten media items. This is notable given that some schools did experience challenges with incorporating the programme into the existing curriculum, and with technology access. The high level of compliance is an indicator of buy-in among teachers and SLT, as it did require scheduling the sessions into the school curriculum, and some teachers reported needing to make trade-offs to accommodate YJA. However, the compliance measure did not capture whether teachers engaged with the intervention delivery or embedded the YJA intervention into their teaching practice. Findings from the case studies indicated that teacher engagement and embedding may be important in realising impacts on pupil writing attainment, and some teachers did not participate in the sessions and some found it challenging to incorporate YJA activities outside of the scheduled sessions. Ensuring that teachers have time to participate in the teaching days and support to incorporate YJA activities and skills in the existing curriculum may be important to support the programme's impact on pupils beyond the mentor's delivery.

Finally, it is important to recognise the intervention was delivered in a real-world setting, and it was being compared against the business-as-usual curriculum. Although the activity of setting up a newsroom is novel, survey data indicated that a high proportion of control schools did engage in journalism-related activities as part of their usual practice. Case study data indicated that YJA differed in that it offered more variety in activities (e.g. television and radio production) and that pupil engagement was enhanced by having an external facilitator for the activities. Nevertheless, an existing usual practice of journalism does set a relatively high bar for YJA to surpass.

Overall, the evaluation does not indicate that YJA has a measurable impact on the primary or secondary outcomes used in this trial. As such, the findings do not support scaling of the programme as delivered. There was evidence that pupils found the programme engaging and teachers perceived benefits, particularly for some pupils who struggle to engage with more traditional approaches. There may be value in revising the intervention and undertaking further research.

Another possible explanation for the discrepancy between teacher-reported outcomes and impact data may be selection bias. Teachers were very positive in their assessment of the YJA intervention and its effect on pupils as writers and the strength and quality of their writing, although this was not supported by the impact evaluation data. The teachers who participated in this intervention may have been self-nominated (if they were not selected by the SLT), displaying a high degree of motivation for the intervention, and it is possible that their perceptions were biased by this, relative to how pupils were actually affected.

Limitations and lessons learnt

Baseline balance checks revealed some imbalance across trial arms, with, on average, higher KS1 scores and higher Ofsted ratings at the schools in the intervention group compared with the control group. The intervention schools in the analysis sample also had a higher proportion of FSM-eligible pupils and a lower proportion of EAL pupils. We acknowledge the potential for this to lead to bias in our point estimates (although, to the extent that it is due to chance in the randomisation this is part of the uncertainty in estimates captured by the confidence intervals). Other things being equal, higher KS1 scores in the intervention group would imply our estimate to be an overestimate of the true impact (we have carried out robustness checking on this point); the same is true for the higher Ofsted rating in the intervention group (we were not able to carry out analogous robustness checking, although correlation between KS1 scores and Ofsted ratings and relatively weak predictive power of Ofsted ratings for pupil attainment over and above this relationship with KS1 scores lead us to doubt this would reveal substantial bias); conversely, the imbalance in terms of FSM pupils might be expected to bias in the other direction. Furthermore, the imbalance in FSM and EAL are both likely to be mitigated by the inclusion of these factors as covariates in our primary analysis model.

Unfortunately, due to issues in accessing the NPD, planned pre-test measures (planned to be phonics screening check scores) were not available at point of analysis. This reduced explanatory power of outcome measure from co-variates included in the primary analysis model compared to our expectations at design stage (although, of course, these were assumptions and so could still have turned out to be too optimistic, even if we did have phonics screening check scores available). This, combined with greater attrition than assumed and a higher intra-cluster correlation of the outcome measure than was anticipated (based on our assumptions and previous EEF guidance), led to a substantial increase in the MDES between the design stage and the analysis stage. This increases the uncertainty around our impact estimates and, hence, decreases the likelihood of finding a statistically significant impact of the intervention for a given effect size. The higher than anticipated ICC of the outcome measure (which we judge to be unlikely to be attributable to the composition of our sample of schools) suggests the need for further preparatory work in understanding this aspect of non-NPD outcome measures ahead of trial design – previous literature on the outcome measure did not cover school-level clustering and neither could our small-scale piloting.

Robust assessment of writing is challenging, particularly during primary schooling. There are few measures available, and none have been used in a similar context. The measure used in this trial, the WAM (Dunsmuir et al., 2015), was a pragmatic choice, which comes with some limitations (e.g. in terms of how relatively new it is); however, existing evidence suggests that it is a valid, consistent and reliable measure. Nevertheless, despite the many dimensions of writing it aims to capture, it may not have been able to capture those proximately affected by this intervention. This does present a limitation to this evaluation, albeit one we feel we have done all we could to address, given the measures available.

Following randomisation, one intervention and eight control schools were lost to follow-up. This led to attrition for this trial. We have to acknowledge that attrition carries with it a risk of bias (Sterne et al., 2019) and one that cannot typically be remedied analytically. In short, those schools and pupils with missing outcomes data could have affected the result from the trial in different ways – meaning that the result is biased away from the 'true' estimate of impact. It was also more difficult to conduct outcome data collection than had been anticipated, resulting in differences in when schools were tested. A robustness check including the date of testing did not reveal any substantial differences in the impact evaluation results.

As discussed in the intervention section, there appear to have been differences in intervention delivery between treated schools. While the evaluation always focuses on a single class, there is also evidence of intervention adaptation likewise to focus on a single class in at least some cases due to concerns about manageability and feasibility. Where the intervention was delivered to an entire year group, these issues of manageability may have resulted in dilution and challenges to the fidelity of the intervention's core aims.

Generally, the data collected from intervention schools as part of the IPE (either via survey or fieldwork visits) only represent the views and experiences of a subset of the larger treatment population. The qualitative findings are therefore not statistically representative, though the use of purposive sampling means that they should provide a good indication of the range and diversity of experiences and attitudes. Additionally, there may be some recall errors in survey responses.

Similarly, the case study approach to the cost evaluation represents the range and diversity of costs encountered in implementation among highly engaged schools. Sampling was done to capture variation of spend among schools with high/low proportions of FSM pupils and per pupil spending, as these were hypothesised to correlate with costs. It is possible that the sampled schools were not representative of typical costs of full implementation of the programme. None of the case study schools in the cost evaluation needed to purchase technological equipment, so the per pupil cost reported does not include what could be a large financial outlay for schools without sufficient technological resources. Spending ranges were provided for other direct costs, so that prospective schools could consider costs they may encounter above and beyond paying for the programme itself.

Future research and publications

The results from the IPE show that pupils enjoyed participating in the YJA intervention and teachers were positive about how it affected their pupils' writing skills and writing self-efficacy. Overall, however, the impact evaluation showed no effect of the YJA intervention on any of the outcome measures. In the logic model, the intervention was

also hypothesised to improve pupil engagement with media and media skills. However, these outcomes were not measured in this trial. Going forward, it would be interesting to understand how participating in interventions such as YJA affects young people's consumption and engagement with news media. We also think there would be value in exploring potential spillovers between YJA activities and attainment in other parts of the curriculum. These are concrete areas where we see the possibility for future research.

An additional overarching report on all five Learning about Culture interventions funded by EEF and the RSA will be published in 2021. This will include the three Key Stage 2 Learning about Culture interventions (Craft of Writing, Power of Pictures, and YJA) and the two Key Stage 1 Learning about Culture interventions (First Thing Music and Speech Bubbles). This report will pool outcome data across the trials for a combined impact evaluation and synthesise IPE results across all interventions.

In addition, we plan to publish several academic journal articles summarising the findings from this and the other trials in the Learning about Culture project, as well as discussing their implications for future research on arts-based learning interventions and methodological issues, including the challenges of measuring children's writing skills at scale.

References

Angrist, J.D., Pischke, J.-S., 2008. *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Bloom, H.S., Richburg-Hayes, L., Black, A.R., 2007. Using covariates to improve precision for studies that randomize schools to evaluate educational interventions. *Educational Evaluation and Policy Analysis* 29, 30–59. https://doi.org/10.3102/0162373707299550

Bruning, R., Dempsey, M., Kauffman, D.F., McKim, C., Zumbrunn, S., 2013. Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology* 105, 25.

Bulus, M., Dong, N., Kelcey, B., Spybrook, J., 2019. PowerUpR: Power analysis tools for multilevel randomized experiments. R package version 1.0.4. https://CRAN.R-project.org/package=PowerUpR

Catterall, J., 2009. *Doing well and doing good by doing art: The effects of education in the visual and performing arts on the achievements and values of young.* Los Angeles/London: Imagination Group/IGroup Books.

Catterall, J.S., 2012. *The arts and achievement in at-risk youth: Findings from four longitudinal studies.* Research Report# 55. National Endowment for the Arts.

Cunningham, S., 2021. *Causal inference: The mixtape.* Yale University Press.

DfE, 2013. English programmes of study: Key Stages 1 and 2. National curriculum in England. Department for Education, London.

DfE, 2016. *Schools, pupils and their characteristics*: *January 2016*, SFR20/2016. Department for Education, London.

DfE, 2017. *Progress in international reading literacy study (PIRLS).* National Report for England. Department for Education, London.

Dunsmuir, S., Kyriacou, M., Batuwitage, S., Hinson, E., Ingram, V., O'Sullivan, S., 2015. An evaluation of the Writing Assessment Measure (WAM) for children's narrative writing. *Assessing Writing* 23, 1–18.

EEF, 2015. Intra-cluster correlation coefficients. Education Endowment Foundation, London.

EEF, 2018. Statistical analysis guidance for EEF evaluations. Education Endowment Foundation, London.

Gerber, A.S., Green, D.P., 2012. *Field experiments: Design, analysis, and interpretation*. WW Norton.

Graham, S., Perin, D., 2007. A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology* 99, 445.

Gutman, L., Schoon, I., 2013. *The impact of non-cognitive skills on outcomes for young people: Literature review*. Education Endowment Foundation.

Hedges, L.V., 1981. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics* 6, 107–128. https://doi.org/10.3102/10769986006002107

Ilie, S., Sutherland, A., Vignoles, A., 2017. Revisiting free school meal eligibility as a proxy for pupil socio-economic deprivation. *British Educational Research Journal* 43, 253–274. https://doi.org/10.1002/berj.3260

Imbens, G.W., Rubin, D.B., 2015. *Causal inference in statistics, social, and biomedical sciences.* Cambridge University Press.

Murphy, V.A., Kyriacou, M., Menon, P., 2013. *Profiling writing challenges in children with English as an Additional Language (EAL).* University of Oxford, Department of Education 14.

Neelands, J., Belfiore, E., Firth, C., Hart, N., Perrin, L., Brock, S., Holdaway, D., Woddis, J., 2015. *Enriching Britain culture, creativity and growth: 2015 report by the Warwick Commission on the future of cultural value*. University of Warwick, Coventry.

Newman, M., Bird, K., Tripney, J., Kalra, N., Kwan, I., Bangpan, M., Vigurs, C., 2010. *Understanding the impact of engagement in culture and sport: A systematic review of the learning impacts for young people.* CASE: The culture and sport evidence programme.

Pajares, F., 1996. Self-efficacy beliefs in academic settings. *Review of Educational Research* 66, 543–578.

Ritchie, J., Lewis, J., Nicholls, C.M., Ormston, R., 2013. *Qualitative research practice: A guide for social science students and researchers*. Sage.

Rubin, D.B., 2004. *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.

Sterne, J.A., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., Cates, C.J., Cheng, H.-Y., Corbett, M.S., Eldridge, S.M., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. *The BMJ* 366.

Strand, S., Malmberg, L., Hall, J., 2015. *English as an Additional Language (EAL) and educational achievement in England: An analysis of the National Pupil Database.* London: Education Endowment Foundation.

Torgerson, C.J., Torgerson, D.J., 2013. *Randomised trials in education: An introductory handbook.* London: Education Endowment Foundation.

UK Government, 2019. Get information about schools. [online].

von Stumm, S., Smith-Woolley, E., Cheesman, R., Pingault, J., Asbury, K., Dale, P.S., Allen, R., Kovas, Y., Plomin, R., 2021. School quality ratings are weak predictors of students' achievement and well-being. *Journal of Child Psychology and Psychiatry* 62, 339–348. https://doi.org/10.1111/jcpp.13276

Weidmann, B., Miratrix, L., 2021. Missing, presumed different: Quantifying the risk of attrition bias in education evaluations. J R Stat Soc Series A rssa.12677. https://doi.org/10.1111/rssa.12677

Appendix A: EEF cost rating

Figure A1: Cost rating

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

Appendix B: Security classification of trial findings

OUTCOME: Writing Assessment Measure (WAM)

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	Risk rating	Comments
Threat 1: Confounding	Low	Some concerns regarding imbalance in school characteristics (e.g., Ofsted rating), discussed in the report.
Threat 2: Concurrent Interventions	Low	Limited information about concurrent interventions.
Threat 3: Experimental effects	Low	IPE for control group schools gave some information about the use of strategies similar to YJA but did not explore any possible changes in practice.
Threat 4: Implementation fidelity	Low	Compliance was generally high in terms of the criteria used in this evaluation, but no information on how the training was implemented.
Threat 5: Missing data	Moderate	There was rather a lot of missing data, with about 24% of the randomisation sample missing. Attrition was considerably higher in the control group. While this is a threat to validity, a padlock has already been dropped for attrition, so no further adjustments are necessary. The lack of NPD and pre-test data also causes some concerns here, though there is little more the authors could have done under the circumstances explained in the report.
Threat 6: Measurement of outcomes	Low/ Moderate	Some concerns about the primary outcome variable, which the authors address at length in the report. Piloting, administration and marking procedures reported provide adequate reassurance.
Threat 7: Selective reporting	Low	The study differs from the original protocol, but the reasons are justified to a satisfactory extent. Selective reporting not apparent.

- **Initial padlock score:** 3 Padlocks – Randomised design with 0.24 MDES at randomisation and 24.5% attrition.
- **Reason for adjustment for threats to validity**: N/A – One moderate and one low/moderate risk, with the direction of likely biases unclear. No further adjustments required.

Final padlock score: initial score adjusted

Appendix C: Effect size estimation

Table C1: Effect size estimation

Outcome	Unadjusted differences in means	Adjusted differences in means	Intervention group		Control group		Pooled SD	Hedges' correction (*J*)
			n (missing)	Outcome SD	*n* (missing)	Outcome SD		
WAM Score (ideas double weighted)	−0.87	−0.66	726 (339)	4.85	915 (170)	4.98	4.92	0.99953
WSEM score	−0.20	0.39	709 (356)	12.12	889 (196)	11.99	12.05	0.99952
Ideation score	−0.06	0.11	709 (356)	4.10	889 (196)	3.86	3.97	0.99952

Further appendices

Appendix D: Additional tables

Table D1: Baseline characteristics of groups as analysed

School-level (categorical)	National-level percentage	Intervention group		Control group		
		N (missing)	%	*N* (missing)	%	
Setting: Urban	87.3	31 (6)	87.0	27 (1)	92.1	
Setting: Rural	12.7	3 (6)	13.0	5 (1)	7.9	
Ofsted: Outstanding	17.1	8 (3)	15.2	4 (1)	22.4	
Ofsted: Good	69.4	26 (3)	71.8	23 (1)	69.7	
Ofsted: RI/Inadequate	13.4	3 (3)	13.0	5 (1)	8.0	
School type: Academy	23.6	14 (2)	41.2	14 (0)	36.9	
School type: Community	41.2	13 (2)	30.6	9 (0)	35.6	
School type: Other	35.2	11 (2)	28.2	10 (0)	27.5	
School-level (continuous)	National-level mean	*N* (missing)	Mean (SD)	*N* (missing)	Mean (SD)	Standardised difference
KS1 average performance	15.9	35 (5)	16.17 (1.29)	30 (3)	16.04 (1.12)	0.107
Pupil-level (categorical)	National-level mean	*n* (missing)	%	*n* (missing)	%	Standardised difference
FSM	30.9	244 (0)	27.5	169 (0)	23.3	
Non-FSM	69.1	643 (0)	72.5	557 (0)	76.7	
EAL	15.3	198 (0)	22.3	181 (0)	24.9	

Non-EAL	84.7	689 (0)	77.7	545 (0)	75.1	

Note. School-level imbalance is calculated applying weights for the size of the school, so that schools which are relatively more important in the pupil-level impact estimation are afforded the same importance in understanding imbalance.

Table D2: Sub-group analyses

Model	Hedges' g (95% CI)	N	p-value of interaction term	p-value of treatment variable
WAM FSM sub-group analysis	−0.25 (−0.51, 0.01)	413	0.56	0.06
WSEM FSM sub-group analysis	−0.10 (−0.20, 0.00)	403	0.21	0.06
Ideation FSM sub-group analysis	−0.16 (−0.33, 0.02)	403	0.03	0.08

Table D3: Full sample summary statistics

Outcome	Mean	SD	ICC	N
WAM Score (ideas scale double weighted)	17.42	4.94	0.29	1641
WSEM-16 Score	63.97	12.08	0.16	1598
Ideation Score	19.88	3.97	0.14	1598

Table D4: Sampling and randomisation inference p-values

Outcome	Effect size	Sampling Inference p-value	Randomisation Inference p-value
WAM Score (ideas scale double weighted)	−0.13	0.16	0.2
WSEM Score	0.03	0.61	0.61
Ideation Score	0.03	0.68	0.69

Table D5: Complier Average Causal Effect (CACE) analysis

Model	Hedges' g (95% CI)	N	First stage F test	Compliance/ treatment correlation	p-value of treatment variable

| Compliance analysis | –0.14 (–0.33, 0.05) | 1613 | F(12, 71) = 2494 | 0.98 | 0.15 |

Table D6: Robustness checks

| Model	Effect size (Hedges' g)	N	Treatment coefficient	p-value of treatment variable
Primary impact evaluation	–0.13	1613	–0.660	0.162
Test date	–0.13	1613	–0.626	0.192
Marker fixed effects	–0.04	1613	–0.204	0.647
School-level KS1	–0.16	1613	–0.797	0.092
MI of primary outcome	–0.13	2108	–0.648	0.098

Table D7: Number of treatment and control schools (N) in each stratum

Randomisation Batch	Stratum	Treatment (N)	Control (N)
1	Low EAL/Low FSM	7	7
1	Low EAL/High FSM	4	3
1	High EAL/Low FSM	3	4
1	High EAL/High FSM	7	6
2	Low EAL/Low FSM	7	7
2	Low EAL/High FSM	3	4
2	High EAL/Low FSM	4	3
2	High EAL/High FSM	6	7

Note: Time of sign up resulted in two randomisation batches. The strata are defined by the intersection of English as an Additional Language (EAL) students (high/low split at sample median within randomisation batch) and proportion of Free School Meals (FSM) students (high/low split at sample median within randomisation batch).
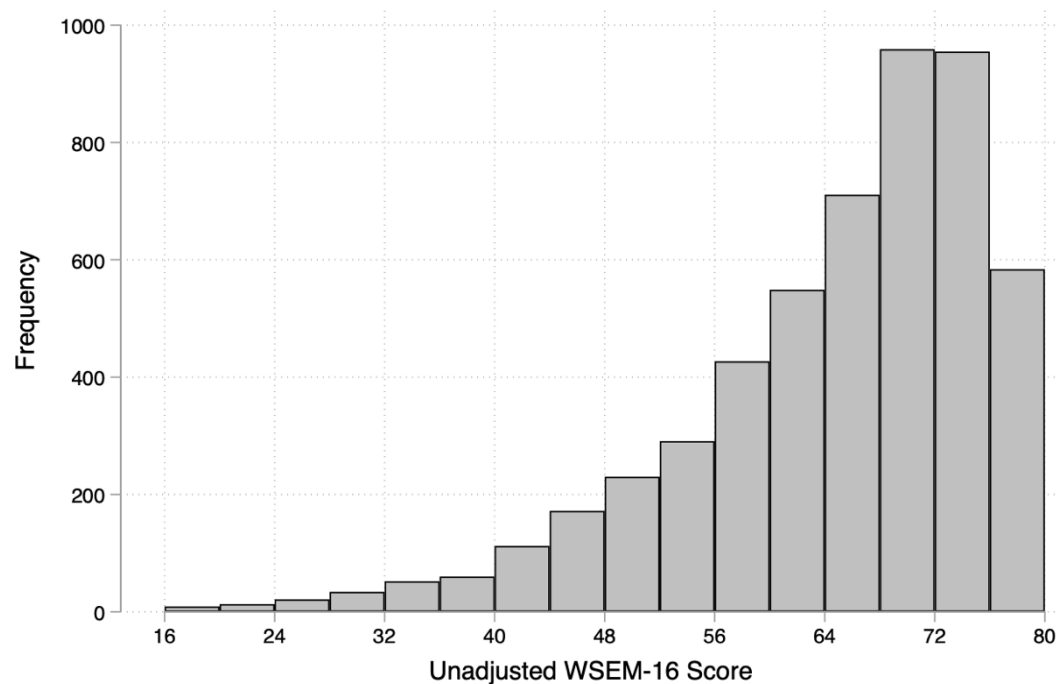
Appendix E: Additional histograms

Figure E1: Histogram of WAM scores pooled across all three KS2 Learning about Culture trials
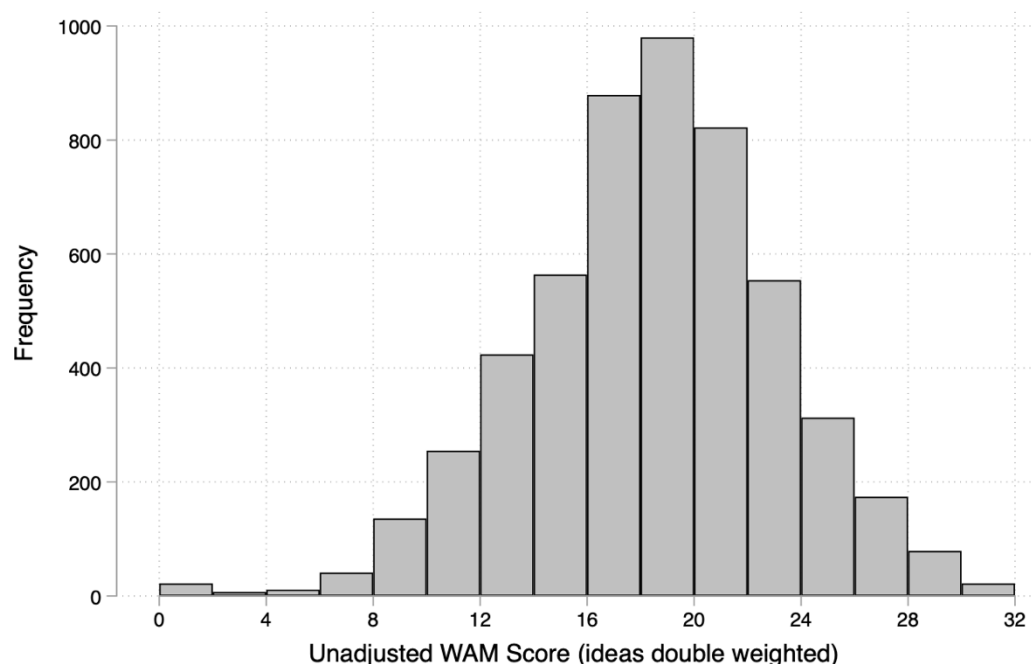
Notes: This histogram contains all WAM outcome data collected for the KS2 Learning about Culture trials (Craft of Writing, Power of Pictures, and Young Journalist Academy). These data were collected as part of a single collection and marking exercise and, as such, we consider it together for the purposes of setting out details of the outcome measures.

Figure E2: Histogram of WSEM scores pooled across all three KS2 Learning about Culture trials
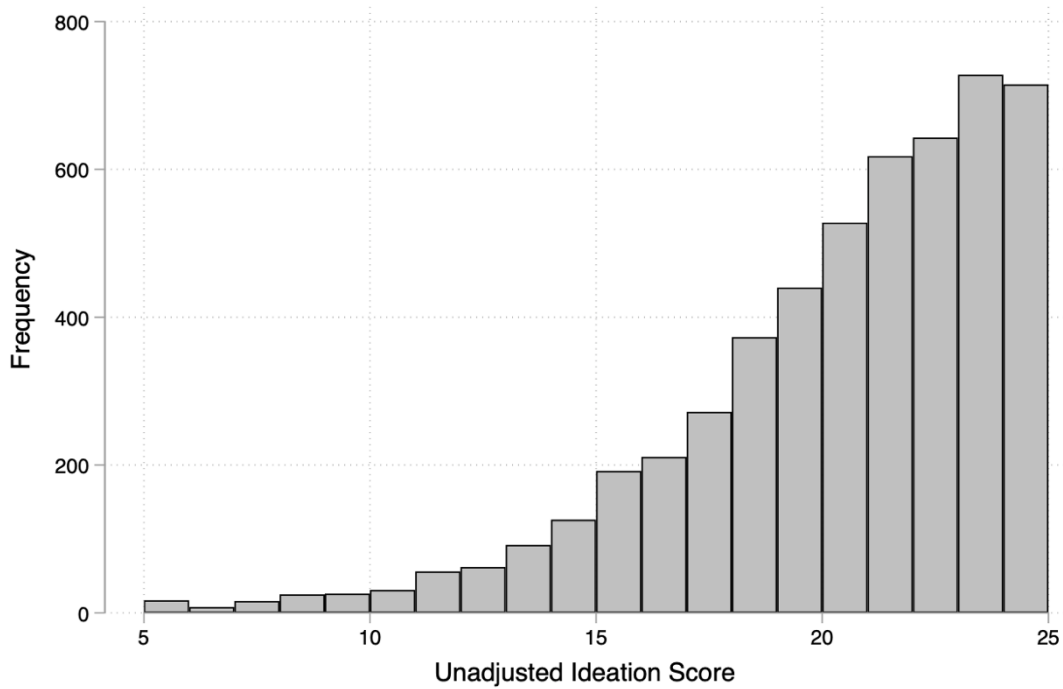

Notes: This histogram contains all WSEMM outcome data collected for the KS2 Learning about Culture trials (Craft of Writing, Power of Pictures, and Young Journalist Academy). These data were collected as part of a single collection and marking exercise and, as such, we consider it together for the purposes of setting out details of the outcome measures.

Figure E3: Histogram of Ideation scores pooled across all three KS2 Learning about Culture trials

Notes: This histogram contains all ideation outcome data collected for the KS2 Learning about Culture trials (Craft of Writing, Power of Pictures, and Young Journalist Academy). These data were collected as part of a single collection and marking exercise and, as such, we consider it together for the purposes of setting out details of the outcome measures.

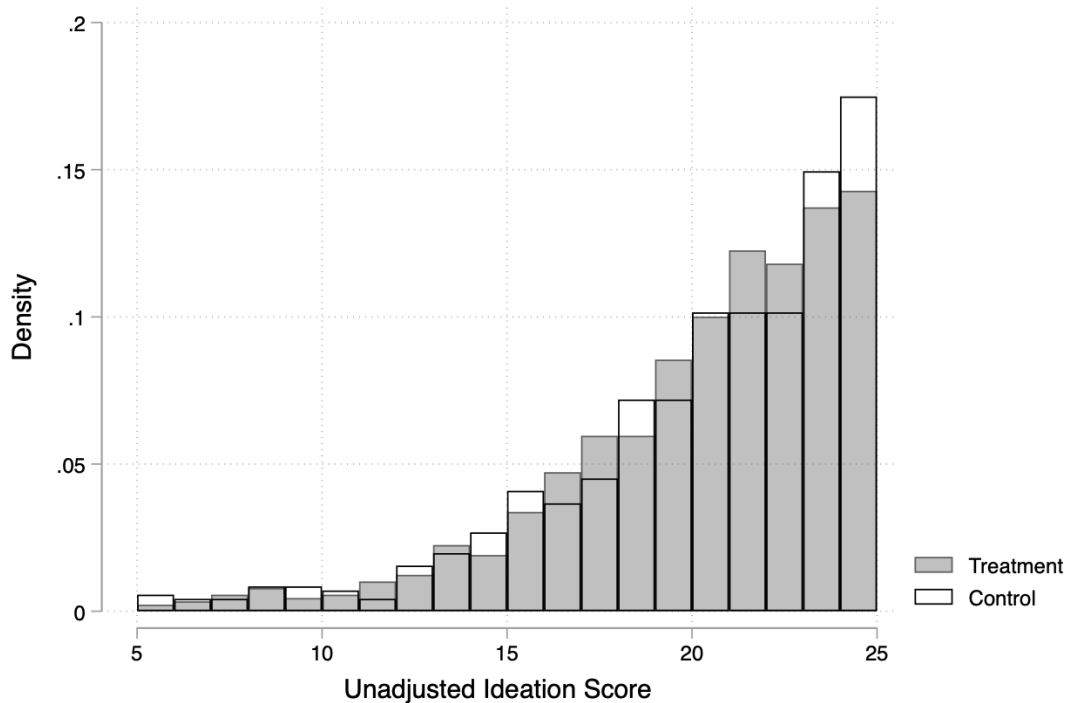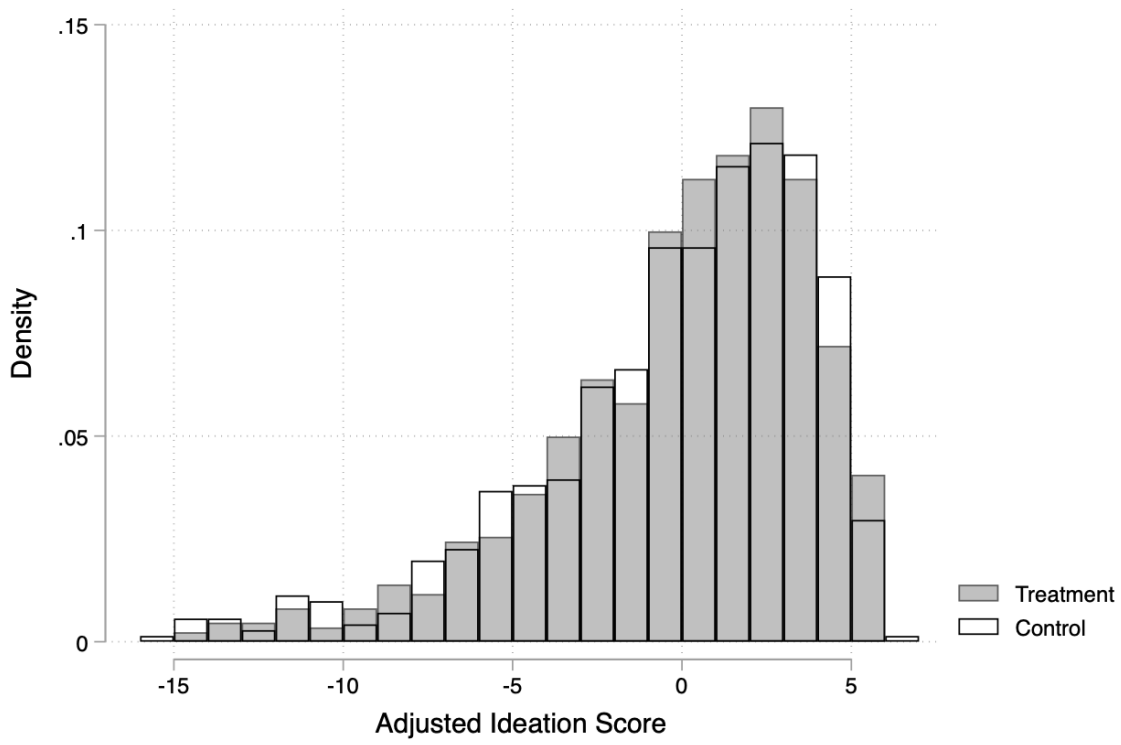**Figure E4: Histogram of unadjusted ideation scores by treatment arm**

Figure E5: Histogram of adjusted ideation scores by treatment arm

Appendix F: Analysis syntax

```
ssc install blindschemes
run `[INSERT FILE PATH HERE]/lac_ks2_nationalmeans.do'
global workd `[INSERT FILE PATH FOR GLOBAL MACRO HERE]'
global datad `[INSERT FILE PATH FOR GLOBAL MACRO HERE]'
cd `${workd}'
clear
use `${datad}/yja_data.dta'
* Create sample variables
cap drop primarysample
gen primarysample = 0
replace primarysample = 1 if wam_score_dw<. & FSM<. & EAL<. & group<. & block<. &
anonschoolid<. & treat<.
count if primarysample == 1
local primarysample_n = r(N)
count if primarysample == 1 & treat==1
local primarysample_treat_n = r(N)
count if primarysample == 1 & treat==0
local primarysample_control_n = r(N)
count if primarysample == 0
local primarysample_miss = r(N)
count if primarysample == 0 & treat==1
local primarysample_treat_miss = r(N)
count if primarysample == 0 & treat==0
local primarysample_control_miss = r(N)
cap drop secondarysample
gen secondarysample = 0
replace secondarysample = 1 if wsem16_score<. & FSM<. & EAL<. & group<. & block<. &
anonschoolid<. & treat<.
count if secondarysample == 1
local secondarysample_n = r(N)
count if secondarysample == 1 & treat==1
local secondarysample_treat_n = r(N)
count if secondarysample == 1 & treat==0
local secondarysample_control_n = r(N)
count if secondarysample == 0
local secondarysample_miss = r(N)
count if secondarysample == 0 & treat==1
local secondarysample_treat_miss = r(N)
count if secondarysample == 0 & treat==0
local secondarysample_control_miss = r(N)
cap drop randomisedsample
```

```
gen randomisedsample = 0

replace randomisedsample = 1 if group<. & block<. & anonschoolid<. & treat<.

count if randomisedsample == 1

local randomisedsample_n = r(N)

count if randomisedsample == 1 & treat==1

local randomisedsample_treat_n = r(N)

count if randomisedsample == 1 & treat==0

local randomisedsample_control_n = r(N)

local primaryattrition_n = `randomisedsample_n' - `primarysample_n'

local primaryattrition_treat_n = `randomisedsample_treat_n' - `primarysample_treat_n'

local primaryattrition_control_n = `randomisedsample_control_n' -
`primarysample_control_n'

local primaryattrition_percent : di %7.1fc 100-
((`primarysample_n'/`randomisedsample_n')*100)

local primaryattrition_treat_p : di %7.1fc 100-
((`primarysample_treat_n'/`randomisedsample_treat_n')*100)

local primaryattrition_control_p : di %7.1fc 100-
((`primarysample_control_n'/`randomisedsample_control_n')*100)

count if wam_score_dw < .

local primarymeasure_n = r(N)

count if wam_score_dw < . & treat==1

local primarymeasure_treat_n = r(N)

count if wam_score_dw < . & treat==0

local primarymeasure_control_n = r(N)

count if wam_score_dw >= .

local primarymeasure_miss = r(N)

count if wam_score_dw >= . & treat==1

local primarymeasure_treat_miss = r(N)

count if wam_score_dw >= . & treat==0

local primarymeasure_control_miss = r(N)

count if wsem16_score < .

local wsem_n = r(N)

count if wsem16_score < . & treat==1

local wsem_treat_n = r(N)

count if wsem16_score < . & treat==0

local wsem_control_n = r(N)

count if wsem16_score >= .

local wsem_miss = r(N)

count if wsem16_score >= . & treat==1

local wsem_treat_miss = r(N)

count if wsem16_score >= . & treat==0

local wsem_control_miss = r(N)

count if ideation_score < .
```

```
local idea_n = r(N)
count if ideation_score < . & treat==1
local idea_treat_n = r(N)
count if ideation_score < . & treat==0
local idea_control_n = r(N)
count if ideation_score >= .
local idea_miss = r(N)
count if ideation_score >= . & treat==1
local idea_treat_miss = r(N)
count if ideation_score >= . & treat==0
local idea_control_miss = r(N)
*Number of treatment and control schools in each block by randomisation group
forval i = 1(1)2{
quietly tab anonschoolid if block == 11 & treat == 0 & group ==`i'
local block11control_schoolcount_`i' = r(r)
quietly tab anonschoolid if block == 11 & treat == 1 & group ==`i'
local block11treat_schoolcount_`i' = r(r)
quietly tab anonschoolid if block == 12 & treat == 0 & group ==`i'
local block12control_schoolcount_`i' = r(r)
quietly tab anonschoolid if block == 12 & treat == 1 & group ==`i'
local block12treat_schoolcount_`i' = r(r)
quietly tab anonschoolid if block == 21 & treat == 0 & group ==`i'
local block21control_schoolcount_`i' = r(r)
quietly tab anonschoolid if block == 21 & treat == 1 & group ==`i'
local block21treat_schoolcount_`i' = r(r)
quietly tab anonschoolid if block == 22 & treat == 0 & group ==`i'
local block22control_schoolcount_`i' = r(r)
quietly tab anonschoolid if block == 22 & treat == 1 & group ==`i'
local block22treat_schoolcount_`i' = r(r)
}
* School-level descriptives
quietly tab anonschoolid
local asrandom_schoolcount = r(r)
quietly tab anonschoolid if treat==0
local asrandom_schoolcount_control = r(r)
quietly tab anonschoolid if treat==1
local asrandom_schoolcount_treat = r(r)
quietly tab anonschoolid if primarysample==1
local wam_schoolcount = r(r)
quietly tab anonschoolid if secondarysample==1
local wsem_schoolcount = r(r)
local idea_schoolcount = r(r)
```

78

```
quietly tab anonschoolid if treat==0 & primarysample==1
local wam_schoolcount_control = r(r)
quietly tab anonschoolid if treat==1 & primarysample==1
local wam_schoolcount_treat = r(r)
quietly tab anonschoolid if treat==0 & secondarysample==1
local wsem_schoolcount_control = r(r)
local idea_schoolcount_control = r(r)
quietly tab anonschoolid if treat==1 & secondarysample==1
local wsem_schoolcount_treat = r(r)
local idea_schoolcount_treat = r(r)
preserve
keep if primarysample==1
gen i=1
collapse (count) count = i, by(anonschoolid)
ameans count
local wam_nhmean : di %7.0fc r(mean_h)
restore
preserve
keep if secondarysample==1
gen i=1
collapse (count) count = i, by(anonschoolid)
ameans count
local wsem_nhmean : di %7.0fc r(mean_h)
local idea_nhmean : di %7.0fc r(mean_h)
restore
sum treat if primarysample==1
local wam_treatprop : di %7.2fc r(mean)
sum treat if secondarysample==1
local wsem_treatprop : di %7.2fc r(mean)
local idea_treatprop : di %7.2fc r(mean)
* School-level descriptives specific to FSM sub-group
preserve
keep if primarysample==1 & FSM==1
gen i=1
collapse (count) count = i, by(anonschoolid)
ameans count
local wam_fsm_nhmean : di %7.0fc r(mean_h)
restore
preserve
keep if secondarysample==1 & FSM==1
gen i=1
collapse (count) count = i, by(anonschoolid)
```

```
ameans count
local wsem_fsm_nhmean : di %7.0fc r(mean_h)
local idea_fsm_nhmean : di %7.0fc r(mean_h)
restore
sum treat if primarysample==1 & FSM==1
local wam_fsm_treatprop : di %7.2fc r(mean)
sum treat if secondarysample==1 & FSM==1
local wsem_fsm_treatprop : di %7.2fc r(mean)
local idea_fsm_treatprop : di %7.2fc r(mean)
* ICC Calculation
xtset anonschoolid
xtreg wam_score_dw
local wam_icc : di %7.2fc e(sigma_u) / (e(sigma_u) + e(sigma_e))
xtreg wsem16_score
local wsem_icc : di %7.2fc e(sigma_u) / (e(sigma_u) + e(sigma_e))
xtreg ideation_score
local idea_icc : di %7.2fc e(sigma_u) / (e(sigma_u) + e(sigma_e))
* ICC Calculation – FSM sub-group
xtset anonschoolid
xtreg wam_score_dw if FSM==1
local wam_fsm_icc : di %7.2fc e(sigma_u) / (e(sigma_u) + e(sigma_e))
xtreg wsem16_score if FSM==1
local wsem_fsm_icc : di %7.2fc e(sigma_u) / (e(sigma_u) + e(sigma_e))
xtreg ideation_score if FSM==1
local idea_fsm_icc : di %7.2fc e(sigma_u) / (e(sigma_u) + e(sigma_e))
* Group variables
cap drop FSMgroup
egen FSMgroup = mean(FSM), by(anonschoolid)
cap drop EALgroup
egen EALgroup = mean(EAL), by(anonschoolid)
* Estimate pre-post test correlations
xtset anonschoolid
xtreg wam_score_dw FSM EAL FSMgroup EALgroup
local wam_r2_pupil : di %7.2fc e(r2_w)
local wam_r2_school : di %7.2fc e(r2_b)
local wam_r2_overall : di %7.2fc e(r2_o)
xtreg wsem16_score FSM EAL FSMgroup EALgroup
local wsem_r2_pupil : di %7.2fc e(r2_w)
local wsem_r2_school : di %7.2fc e(r2_b)
local wsem_r2_overall : di %7.2fc e(r2_o)
xtreg ideation_score FSM EAL FSMgroup EALgroup
local idea_r2_pupil : di %7.2fc e(r2_w)
```

```
local idea_r2_school : di %7.2fc e(r2_b)
local idea_r2_overall : di %7.2fc e(r2_o)
* Estimate pre-post test correlations - FSM sub-group
xtset anonschoolid
xtreg wam_score_dw EAL FSMgroup EALgroup if FSM==1
local wam_fsm_r2_pupil : di %7.2fc e(r2_w)
local wam_fsm_r2_school : di %7.2fc e(r2_b)
local wam_fsm_r2_overall : di %7.2fc e(r2_o)
xtreg wsem16_score EAL FSMgroup EALgroup if FSM==1
local wsem_fsm_r2_pupil : di %7.2fc e(r2_w)
local wsem_fsm_r2_school : di %7.2fc e(r2_b)
local wsem_fsm_r2_overall : di %7.2fc e(r2_o)
xtreg ideation_score EAL FSMgroup EALgroup if FSM==1
local idea_fsm_r2_pupil : di %7.2fc e(r2_w)
local idea_fsm_r2_school : di %7.2fc e(r2_b)
local idea_fsm_r2_overall : di %7.2fc e(r2_o)
** Imbalance at baseline - as randomised
* FSM
sum FSM if treat==0 & randomisedsample==1
local fsm_controlpercent_in_rsample : di %7.1fc r(mean)*100
local fsm_controlpercent_out_rsample : di %7.1fc (1-r(mean))*100
local fsm_controlmean_rsample : di %7.2fc r(mean)
local fsm_controlsd_rsample : di %7.2fc r(sd)
local fsm_controln_rsample = r(N)
count if FSM==0 & treat==0 & randomisedsample==1
local fsm_controln_out_rsample = r(N)
count if FSM==1 & treat==0 & randomisedsample==1
local fsm_controln_in_rsample = r(N)
count if FSM>=. & treat==0 & randomisedsample==1
local fsm_controlmiss_rsample = r(N)
sum FSM if treat==1 & randomisedsample==1
local fsm_treatpercent_in_rsample : di %7.1fc r(mean)*100
local fsm_treatpercent_out_rsample : di %7.1fc (1-r(mean))*100
local fsm_treatmean_rsample : di %7.2fc r(mean)
local fsm_treatsd_rsample : di %7.2fc r(sd)
local fsm_treatn_rsample = r(N)
count if FSM==0 & treat==1 & randomisedsample==1
local fsm_treatn_out_rsample = r(N)
count if FSM==1 & treat==1 & randomisedsample==1
local fsm_treatn_in_rsample = r(N)
count if FSM>=. & treat==1 & randomisedsample==1
local fsm_treatmiss_rsample = r(N)
```

81

```
esizei `fsm_controln_rsample' `fsm_controlmean_rsample' `fsm_controlsd_rsample' ///
      `fsm_treatn_rsample' `fsm_treatmean_rsample' `fsm_treatsd_rsample', cohensd
local fsm_in_stddiff_rsample : di %7.3fc r(d)
local fsm_out_stddiff_rsample : di %7.3fc -r(d)
* EAL
sum EAL if treat==0 & randomisedsample==1
local eal_controlpercent_in_rsample : di %7.1fc r(mean)*100
local eal_controlpercent_out_rsample : di %7.1fc (1-r(mean))*100
local eal_controlmean_rsample : di %7.2fc r(mean)
local eal_controlsd_rsample : di %7.2fc r(sd)
local eal_controln_rsample = r(N)
count if EAL==0 & treat==0 & randomisedsample==1
local eal_controln_out_rsample = r(N)
count if EAL==1 & treat==0 & randomisedsample==1
local eal_controln_in_rsample = r(N)
count if EAL>=. & treat==0 & randomisedsample==1
local eal_controlmiss_rsample = r(N)
sum EAL if treat==1 & randomisedsample==1
local eal_treatpercent_in_rsample : di %7.1fc r(mean)*100
local eal_treatpercent_out_rsample : di %7.1fc (1-r(mean))*100
local eal_treatmean_rsample : di %7.2fc r(mean)
local eal_treatsd_rsample : di %7.2fc r(sd)
local eal_treatn_rsample = r(N)
count if EAL==0 & treat==1 & randomisedsample==1
local eal_treatn_out_rsample = r(N)
count if EAL==1 & treat==1 & randomisedsample==1
local eal_treatn_in_rsample = r(N)
count if EAL>=. & treat==1 & randomisedsample==1
local eal_treatmiss_rsample = r(N)
esizei `eal_controln_rsample' `eal_controlmean_rsample' `eal_controlsd_rsample' ///
      `eal_treatn_rsample' `eal_treatmean_rsample' `eal_treatsd_rsample', cohensd
local eal_in_stddiff_rsample : di %7.3fc r(d)
local eal_out_stddiff_rsample : di %7.3fc -r(d)
** Imbalance at baseline - as analysed
* FSM
sum FSM if treat==0 & primarysample==1
local fsm_controlpercent_in_psample : di %7.1fc r(mean)*100
local fsm_controlpercent_out_psample : di %7.1fc (1-r(mean))*100
local fsm_controlmean_psample : di %7.2fc r(mean)
local fsm_controlsd_psample : di %7.2fc r(sd)
local fsm_controln_psample = r(N)
count if FSM==0 & treat==0 & primarysample==1
```

```
local fsm_controln_out_psample = r(N)
count if FSM==1 & treat==0 & primarysample==1
local fsm_controln_in_psample = r(N)
count if FSM>=. & treat==0 & primarysample==1
local fsm_controlmiss_psample = r(N)
sum FSM if treat==1 & primarysample==1
local fsm_treatpercent_in_psample : di %7.1fc r(mean)*100
local fsm_treatpercent_out_psample : di %7.1fc (1-r(mean))*100
local fsm_treatmean_psample : di %7.2fc r(mean)
local fsm_treatsd_psample : di %7.2fc r(sd)
local fsm_treatn_psample = r(N)
count if FSM==0 & treat==1 & primarysample==1
local fsm_treatn_out_psample = r(N)
count if FSM==1 & treat==1 & primarysample==1
local fsm_treatn_in_psample = r(N)
count if FSM>=. & treat==0 & primarysample==1
local fsm_treatmiss_psample = r(N)
esizei `fsm_controln_psample' `fsm_controlmean_psample' `fsm_controlsd_psample' ///
      `fsm_treatn_psample' `fsm_treatmean_psample' `fsm_treatsd_psample', cohensd
local fsm_in_stddiff_psample : di %7.3fc r(d)
local fsm_out_stddiff_psample : di %7.3fc -r(d)
* EAL
sum EAL if treat==0 & primarysample==1
local eal_controlpercent_in_psample : di %7.1fc r(mean)*100
local eal_controlpercent_out_psample : di %7.1fc (1-r(mean))*100
local eal_controlmean_psample : di %7.2fc r(mean)
local eal_controlsd_psample : di %7.2fc r(sd)
local eal_controln_psample = r(N)
count if EAL==0 & treat==0 & primarysample==1
local eal_controln_out_psample = r(N)
count if EAL==1 & treat==0 & primarysample==1
local eal_controln_in_psample = r(N)
count if EAL>=. & treat==0 & primarysample==1
local eal_controlmiss_psample = r(N)
sum EAL if treat==1 & primarysample==1
local eal_treatpercent_in_psample : di %7.1fc r(mean)*100
local eal_treatpercent_out_psample : di %7.1fc (1-r(mean))*100
local eal_treatmean_psample : di %7.2fc r(mean)
local eal_treatsd_psample : di %7.2fc r(sd)
local eal_treatn_psample = r(N)
count if EAL==0 & treat==1 & primarysample==1
local eal_treatn_out_psample = r(N)
```

```
count if EAL==1 & treat==1 & primarysample==1

local eal_treatn_in_psample = r(N)

count if EAL>=. & treat==0 & primarysample==1

local eal_treatmiss_psample = r(N)

esizei `eal_controln_psample' `eal_controlmean_psample' `eal_controlsd_psample' ///
    `eal_treatn_psample' `eal_treatmean_psample' `eal_treatsd_psample', cohensd

local eal_in_stddiff_psample : di %7.3fc r(d)

local eal_out_stddiff_psample : di %7.3fc -r(d)

* Visualisation of outcome variables

graph twoway (kdensity wam_score_dw), scheme(plotplain) xtitle('Unadjusted WAM Score
(ideas double weighted)') ytitle('Density')

graph export output/wam_kdensity.png, replace

graph twoway (kdensity wsem16_score), scheme(plotplain) xtitle('Unadjusted WSEM-16
Score') ytitle('Density')

graph export output/wsem_kdensity.png, replace

graph twoway (kdensity ideation_score), scheme(plotplain) xtitle('Unadjusted Ideation
Score') ytitle('Density')

graph export output/ideation_kdensity.png, replace

* Descriptive statistics for outcome variables

sum wam_score_dw

local wam_mean : di %7.2fc r(mean)

local wam_sd : di %7.2fc r(sd)

local wam_n = r(N)

sum wsem16_score

local wsem_mean : di %7.2fc r(mean)

local wsem_sd : di %7.2fc r(sd)

local wsem_n = r(N)

sum ideation_score

local idea_mean : di %7.2fc r(mean)

local idea_sd : di %7.2fc r(sd)

local idea_n = r(N)

* Primary analysis

regress wam_score_dw i.treat FSM EAL FSMgroup EALgroup group#block, vce(cluster
anonschoolid)

estimates store primary

local wam_treatdiff : di %7.2fc _b[1.treat]

local wam_treatdiff_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]

local wam_treatdiff_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]

local wam_treatdiff_p : di %7.2fc ttail(e(df_r), abs(_b[1.treat]/_se[1.treat]))*2

margins treat, post

local wam_controlmean : di %7.2fc _b[0.treat]

local wam_controlmean_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]

local wam_controlmean_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]
```

```
local wam_treatmean : di %7.2fc _b[1.treat]
local wam_treatmean_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local wam_treatmean_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
* Unadjusted version
regress wam_score_dw i.treat, vce(cluster anonschoolid)
local wam_treatdiff_unadj : di %7.2fc _b[1.treat]
local wam_treatdiff_unadj_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local wam_treatdiff_unadj_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
local wam_treatdiff_unadj_p : di %7.2fc ttail(e(df_r), abs(_b[1.treat]/_se[1.treat]))*2
margins treat, post
local wam_controlmean_unadj : di %7.2fc _b[0.treat]
local wam_controlmean_unadj_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]
local wam_controlmean_unadj_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]
local wam_treatmean_unadj : di %7.2fc _b[1.treat]
local wam_treatmean_unadj_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local wam_treatmean_unadj_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
* Graphical version of primary analysis
regress wam_score_dw FSM EAL FSMgroup EALgroup group#block, vce(cluster anonschoolid)
cap drop wam_analysis_resid
predict wam_analysis_resid, resid
graph twoway ///
      (kdensity wam_score_dw if treat==1) ///
      (kdensity wam_score_dw if treat==0) ///
      , scheme(plotplain) legend(order(1 'Treatment' 2 'Control')) xtitle('Unadjusted
WAM Score') ytitle('Density')
graph export output/wam_kdensity_treat_uncond.png, replace
graph twoway ///
      (kdensity wam_analysis_resid if treat==1) ///
      (kdensity wam_analysis_resid if treat==0) ///
      , scheme(plotplain) legend(order(1 'Treatment' 2 'Control')) xtitle('Adjusted WAM
Score') ytitle('Density')
graph export output/wam_kdensity_treat_cond.png, replace
* Secondary analysis - WSEM
regress wsem16_score i.treat FSM EAL FSMgroup EALgroup group#block, vce(cluster
anonschoolid)
local wsem_treatdiff : di %7.2fc _b[1.treat]
local wsem_treatdiff_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local wsem_treatdiff_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
local wsem_treatdiff_p : di %7.2fc ttail(e(df_r), abs(_b[1.treat]/_se[1.treat]))*2
margins treat, post
local wsem_controlmean : di %7.2fc _b[0.treat]
local wsem_controlmean_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]
local wsem_controlmean_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]
```

```
local wsem_treatmean : di %7.2fc _b[1.treat]
local wsem_treatmean_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local wsem_treatmean_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
* Unadjusted version
regress wsem16_score i.treat, vce(cluster anonschoolid)
local wsem_treatdiff_unadj : di %7.2fc _b[1.treat]
local wsem_treatdiff_unadj_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local wsem_treatdiff_unadj_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
local wsem_treatdiff_unadj_p : di %7.2fc ttail(e(df_r),
abs(_b[1.treat]/_se[1.treat]))*2
margins treat, post
local wsem_controlmean_unadj : di %7.2fc _b[0.treat]
local wsem_controlmean_unadj_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]
local wsem_controlmean_unadj_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]
local wsem_treatmean_unadj : di %7.2fc _b[1.treat]
local wsem_treatmean_unadj_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local wsem_treatmean_unadj_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
* Graphical version of secondary analysis - WSEM
regress wsem16_score FSM EAL FSMgroup EALgroup group#block, vce(cluster anonschoolid)
cap drop wsem_analysis_resid
predict wsem_analysis_resid, resid
graph twoway ///
      (kdensity wsem16_score if treat==1) ///
      (kdensity wsem16_score if treat==0) ///
      , scheme(plotplain) legend(order(1 'Treatment' 2 'Control')) xtitle('Unadjusted
WSEM Score') ytitle('Density')
graph export output/wsem_kdensity_treat_uncond.png, replace
graph twoway ///
      (kdensity wsem_analysis_resid if treat==1) ///
      (kdensity wsem_analysis_resid if treat==0) ///
      , scheme(plotplain) legend(order(1 'Treatment' 2 'Control')) xtitle('Adjusted
WSEM Score') ytitle('Density')
graph export output/wsem_kdensity_treat_cond.png, replace
* Secondary analysis - Ideation
regress ideation_score i.treat FSM EAL FSMgroup EALgroup group#block, vce(cluster
anonschoolid)
local idea_treatdiff : di %7.2fc _b[1.treat]
local idea_treatdiff_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local idea_treatdiff_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
local idea_treatdiff_p : di %7.2fc ttail(e(df_r), abs(_b[1.treat]/_se[1.treat]))*2
margins treat, post
local idea_controlmean : di %7.2fc _b[0.treat]
local idea_controlmean_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]
```

```
local idea_controlmean_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]
local idea_treatmean : di %7.2fc _b[1.treat]
local idea_treatmean_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local idea_treatmean_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
regress ideation_score i.treat, vce(cluster anonschoolid)
local idea_treatdiff_unadj : di %7.2fc _b[1.treat]
local idea_treatdiff_unadj_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local idea_treatdiff_unadj_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
local idea_treatdiff_unadj_p : di %7.2fc ttail(e(df_r),
abs(_b[1.treat]/_se[1.treat]))*2
margins treat, post
local idea_controlmean_unadj : di %7.2fc _b[0.treat]
local idea_controlmean_unadj_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]
local idea_controlmean_unadj_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]
local idea_treatmean_unadj : di %7.2fc _b[1.treat]
local idea_treatmean_unadj_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]
local idea_treatmean_unadj_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]
* Graphical version of secondary analysis - ideation
regress ideation_score FSM EAL FSMgroup EALgroup group#block, vce(cluster anonschoolid)
cap drop ideation_analysis_resid
predict ideation_analysis_resid, resid
graph twoway ///
     (kdensity ideation_score if treat==1) ///
     (kdensity ideation_score if treat==0) ///
     , scheme(plotplain) legend(order(1 'Treatment' 2 'Control')) xtitle('Unadjusted
Ideation Score') ytitle('Density')
graph export output/ideation_kdensity_treat_uncond.png, replace
graph twoway ///
     (kdensity ideation_analysis_resid if treat==1) ///
     (kdensity ideation_analysis_resid if treat==0) ///
     , scheme(plotplain) legend(order(1 'Treatment' 2 'Control')) xtitle('Adjusted
Ideation Score') ytitle('Density')
graph export output/ideation_kdensity_treat_cond.png, replace
* Effect size calculations - primary analysis
count if primarysample==1 & treat==0
local wam_controln = r(N)
count if primarysample==1 & treat==1
local wam_treatn = r(N)
quietly summarize wam_score_dw if primarysample==1 & treat==0
local wam_controlsd : di %7.2fc r(sd)
quietly summarize wam_score_dw if primarysample==1 & treat==1
local wam_treatsd : di %7.2fc r(sd)
```

```
local wam_sdpooled = sqrt(((`wam_controln'-1)*(`wam_controlsd'^2) + (`wam_treatn'-
1)*(`wam_treatsd'^2))/(`wam_controln'+`wam_treatn'-2))

local wam_sdpooled : di %7.2fc `wam_sdpooled'

local wam_j = (exp(lngamma((`wam_treatn'+`wam_controln'+2)/2)))) /
(((`wam_treatn'+`wam_controln'+2)/2)^.5 * exp(lngamma((`wam_treatn'+`wam_controln'+2-
1)/2)))

if `wam_j' == . {
     local wam_j = 1-(3/((4*(`wam_treatn'+`wam_controln'))-9))
     }

local wam_g : di %7.2fc `wam_j'*(`wam_treatdiff'/`wam_sdpooled')

local wam_g_upperci : di %7.2fc `wam_j'*(`wam_treatdiff_upperci'/`wam_sdpooled')

local wam_g_lowerci : di %7.2fc `wam_j'*(`wam_treatdiff_lowerci'/`wam_sdpooled')

local wam_j : di %7.5fc `wam_j'

* Effect size calculations - secondary analysis - WSEM

count if wsem16_score<. & secondarysample==1 & treat==0

local wsem_controln = r(N)

count if wsem16_score<. & secondarysample==1 & treat==1

local wsem_treatn = r(N)

quietly summarize wsem16_score if secondarysample==1 & treat==0

local wsem_controlsd : di %7.2fc r(sd)

quietly summarize wsem16_score if secondarysample==1 & treat==1

local wsem_treatsd : di %7.2fc r(sd)

local wsem_sdpooled = sqrt(((`wsem_controln'-1)*(`wsem_controlsd'^2) + (`wsem_treatn'-
1)*(`wsem_treatsd'^2))/(`wsem_controln'+`wsem_treatn'-2))

local wsem_sdpooled : di %7.2fc `wsem_sdpooled'

local wsem_j = (exp(lngamma((`wsem_treatn'+`wsem_controln'+2)/2)))) /
(((`wsem_treatn'+`wsem_controln'+2)/2)^.5 *
exp(lngamma((`wsem_treatn'+`wsem_controln'+2-1)/2)))

if `wsem_j' == . {
     local wsem_j = 1-(3/((4*(`wsem_treatn'+`wsem_controln'))-9))
     }

local wsem_g : di %7.2fc `wsem_j'*(`wsem_treatdiff'/`wsem_sdpooled')

local wsem_g_upperci : di %7.2fc `wsem_j'*(`wsem_treatdiff_upperci'/`wsem_sdpooled')

local wsem_g_lowerci : di %7.2fc `wsem_j'*(`wsem_treatdiff_lowerci'/`wsem_sdpooled')

local wsem_j : di %7.5fc `wsem_j'

* Effect size calculations - secondary analysis - Ideation

count if ideation_score<. & secondarysample==1 & treat==0

local idea_controln = r(N)

count if ideation_score<. & secondarysample==1 & treat==1

local idea_treatn = r(N)

quietly summarize ideation_score if secondarysample==1 & treat==0

local idea_controlsd : di %7.2fc r(sd)

quietly summarize ideation_score if secondarysample==1 & treat==1

local idea_treatsd : di %7.2fc r(sd)
```

```
local idea_sdpooled = sqrt((((`idea_controln'-1)*(`idea_controlsd'^2) + (`idea_treatn'-
1)*(`idea_treatsd'^2))/(`idea_controln'+`idea_treatn'-2))

local idea_sdpooled : di %7.2fc `idea_sdpooled'

local idea_j = (exp(lngamma((`idea_treatn'+`idea_controln'+2)/2))) /
(((`idea_treatn'+`idea_controln'+2)/2)^.5 *
exp(lngamma((`idea_treatn'+`idea_controln'+2-1)/2)))

if `idea_j' == . {

        local idea_j = 1-(3/((4*(`idea_treatn'+`idea_controln'))-9))

        }

local idea_g : di %7.2fc `idea_j'*(`idea_treatdiff'/`idea_sdpooled')

local idea_g_upperci : di %7.2fc `idea_j'*(`idea_treatdiff_upperci'/`idea_sdpooled')

local idea_g_lowerci : di %7.2fc `idea_j'*(`idea_treatdiff_lowerci'/`idea_sdpooled')

local idea_j : di %7.5fc `idea_j'

* Complier analysis

corr treat comply

local complier_r : di %7.2fc r(rho)

regress comply FSM EAL FSMgroup EALgroup group#block treat if primarysample==1,
vce(cluster anonschoolid)

local complier_F : di %7.0f e(F)

local complier_df_m = e(df_m)

local complier_df_r = e(df_r)

ivregress 2sls wam_score_dw FSM EAL FSMgroup EALgroup group#block (i.comply = treat) if
primarysample==1, vce(cluster anonschoolid)

local complier_treatdiff : di %7.2fc _b[1.comply]

local complier_treatdiff_upperci : di %7.2fc _b[1.comply] + 1.96*_se[1.comply]

local complier_treatdiff_lowerci : di %7.2fc _b[1.comply] - 1.96*_se[1.comply]

local complier_treatdiff_p : di %7.2fc normal(-abs(_b[1.comply]/_se[1.comply]))*2

margins comply, post

local complier_controlmean : di %7.2fc _b[0.comply]

local complier_controlmean_upperci : di %7.2fc _b[0.comply] + 1.96*_se[0.comply]

local complier_controlmean_lowerci : di %7.2fc _b[0.comply] - 1.96*_se[0.comply]

local complier_treatmean : di %7.2fc _b[1.comply]

local complier_treatmean_upperci : di %7.2fc _b[1.comply] + 1.96*_se[1.comply]

local complier_treatmean_lowerci : di %7.2fc _b[1.comply] - 1.96*_se[1.comply]

* Effect size calculations - complier analysis

count if primarysample==1 & comply<. & treat==0

local complier_controln = r(N)

count if primarysample==1 & comply<. & treat==1

local complier_treatn = r(N)

count if primarysample==1 & comply<.

local complier_n = r(N)

quietly summarize wam_score_dw if primarysample==1 & comply<. & treat==0

local complier_controlsd : di %7.2fc r(sd)

quietly summarize wam_score_dw if primarysample==1 & comply<. & treat==1
```

89

```
local complier_treatsd : di %7.2fc r(sd)

local complier_sdpooled = sqrt(((`complier_controln'-1)*(`complier_controlsd'^2) +
(`complier_treatn'-1)*(`complier_treatsd'^2))/(`complier_controln'+`complier_treatn'-
2))

local complier_sdpooled : di %7.2fc `complier_sdpooled'

local complier_j = (exp(lngamma((`complier_treatn'+`complier_controln'+2)/2))) /
((((`complier_treatn'+`complier_controln'+2)/2)^.5 *
exp(lngamma((`complier_treatn'+`complier_controln'+2-1)/2)))

if `complier_j' == . {

        local complier_j = 1-(3/((4*(`complier_treatn'+`complier_controln'))-9))

        }

local complier_g : di %7.2fc `complier_j'*(`complier_treatdiff'/`complier_sdpooled')

local complier_g_upperci : di %7.2fc
`complier_j'*(`complier_treatdiff_upperci'/`complier_sdpooled')

local complier_g_lowerci : di %7.2fc
`complier_j'*(`complier_treatdiff_lowerci'/`complier_sdpooled')

* FSM check for interaction then run for EVERFSM_6_P sub-group - primary analysis

regress wam_score_dw i.treat i.FSM treat#FSM EAL FSMgroup EALgroup group#block,
vce(cluster anonschoolid)

local fsmwam_interaction_p : di %7.2fc ttail(e(df_r),
abs(_b[1.treat#1.FSM]/_se[1.treat#1.FSM]))*2

regress wam_score_dw i.treat EAL FSMgroup EALgroup group#block if FSM==1, vce(cluster
anonschoolid)

local fsmwam_treatdiff : di %7.2fc _b[1.treat]

local fsmwam_treatdiff_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]

local fsmwam_treatdiff_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]

local fsmwam_treatdiff_p : di %7.2fc ttail(e(df_r), abs(_b[1.treat]/_se[1.treat]))*2

margins treat, post

local fsmwam_controlmean : di %7.2fc _b[0.treat]

local fsmwam_controlmean_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]

local fsmwam_controlmean_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]

local fsmwam_treatmean : di %7.2fc _b[1.treat]

local fsmwam_treatmean_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]

local fsmwam_treatmean_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]

* FSM check for interaction then run for EVERFSM_6_P sub-group - secondary analysis
(WSEM)

regress wsem16_score i.treat i.FSM treat#FSM EAL FSMgroup EALgroup group#block,
vce(cluster anonschoolid)

local fsmwsem_interaction_p : di %7.2fc ttail(e(df_r),
abs(_b[1.treat#1.FSM]/_se[1.treat#1.FSM]))*2

regress wam_score_dw i.treat EAL FSMgroup EALgroup group#block if FSM==1, vce(cluster
anonschoolid)

local fsmwsem_treatdiff : di %7.2fc _b[1.treat]

local fsmwsem_treatdiff_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]

local fsmwsem_treatdiff_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]

local fsmwsem_treatdiff_p : di %7.2fc ttail(e(df_r), abs(_b[1.treat]/_se[1.treat]))*2
```

```
margins treat, post

local fsmwsem_controlmean : di %7.2fc _b[0.treat]

local fsmwsem_controlmean_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]

local fsmwsem_controlmean_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]

local fsmwsem_treatmean : di %7.2fc _b[1.treat]

local fsmwsem_treatmean_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]

local fsmwsem_treatmean_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]

* FSM check for interaction then run for EVERFSM_6_P sub-group - secondary analysis
(Ideation)

regress ideation_score i.treat i.FSM treat#FSM EAL FSMgroup EALgroup group#block,
vce(cluster anonschoolid)

local fsmidea_interaction_p : di %7.2fc ttail(e(df_r),
abs(_b[1.treat#1.FSM]/_se[1.treat#1.FSM]))*2

regress ideation_score i.treat EAL FSMgroup EALgroup group#block if FSM==1, vce(cluster
anonschoolid)

local fsmidea_treatdiff : di %7.2fc _b[1.treat]

local fsmidea_treatdiff_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]

local fsmidea_treatdiff_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]

local fsmidea_treatdiff_p : di %7.2fc ttail(e(df_r), abs(_b[1.treat]/_se[1.treat]))*2

margins treat, post

local fsmidea_controlmean : di %7.2fc _b[0.treat]

local fsmidea_controlmean_upperci : di %7.2fc _b[0.treat] + 1.96*_se[0.treat]

local fsmidea_controlmean_lowerci : di %7.2fc _b[0.treat] - 1.96*_se[0.treat]

local fsmidea_treatmean : di %7.2fc _b[1.treat]

local fsmidea_treatmean_upperci : di %7.2fc _b[1.treat] + 1.96*_se[1.treat]

local fsmidea_treatmean_lowerci : di %7.2fc _b[1.treat] - 1.96*_se[1.treat]

* Effect size calculations - FSM sub-group analysis

count if primarysample==1 & FSM==1 & treat==0

local primarysample_fsm_control_n = r(N)

count if primarysample==1 & FSM==1 & treat==1

local primarysample_fsm_treat_n = r(N)

count if primarysample==1 & FSM==1

local primarysample_fsm_n = r(N)

count if secondarysample==1 & FSM==1 & treat==0

local secondarysample_fsm_control_n = r(N)

count if secondarysample==1 & FSM==1 & treat==1

local secondarysample_fsm_treat_n = r(N)

count if secondarysample==1 & FSM==1

local secondarysample_fsm_n = r(N)

* Primary

quietly summarize wam_score_dw if primarysample==1 & FSM==1 & treat==0

local fsmwam_controlsd : di %7.2fc r(sd)

quietly summarize wam_score_dw if primarysample==1 & FSM==1 & treat==1
```

```
local fsmwam_treatsd : di %7.2fc r(sd)

local fsmwam_sdpooled = sqrt(((`primarysample_fsm_control_n'-1)*(`fsmwam_controlsd'^2)
+ (`primarysample_fsm_treat_n'-
1)*(`fsmwam_treatsd'^2))/(`primarysample_fsm_control_n'+`primarysample_fsm_treat_n'-2))

local fsmwam_sdpooled : di %7.2fc `fsmwam_sdpooled'

local fsmwam_j =
(exp(lngamma((`primarysample_fsm_treat_n'+`primarysample_fsm_control_n'+2)/2)))) /
(((`primarysample_fsm_treat_n'+`primarysample_fsm_control_n'+2)/2)^.5 *
exp(lngamma((`primarysample_fsm_treat_n'+`primarysample_fsm_control_n'+2-1)/2)))

if `fsmwam_j' == . {

    local fsmwam_j = 1-
(3/((4*(`primarysample_fsm_treat_n'+`primarysample_fsm_control_n'))-9))

    }

local fsmwam_g : di %7.2fc `fsmwam_j'*(`fsmwam_treatdiff'/`fsmwam_sdpooled')

local fsmwam_g_upperci : di %7.2fc
`fsmwam_j'*(`fsmwam_treatdiff_upperci'/`fsmwam_sdpooled')

local fsmwam_g_lowerci : di %7.2fc
`fsmwam_j'*(`fsmwam_treatdiff_lowerci'/`fsmwam_sdpooled')

* Secondary (WSEM)

quietly summarize wsem16_score if secondarysample==1 & FSM==1 & treat==0

local fsmwsem_controlsd : di %7.2fc r(sd)

quietly summarize wsem16_score if secondarysample==1 & FSM==1 & treat==1

local fsmwsem_treatsd : di %7.2fc r(sd)

local fsmwsem_sdpooled = sqrt(((`secondarysample_fsm_control_n'-
1)*(`fsmwsem_controlsd'^2) + (`secondarysample_fsm_treat_n'-
1)*(`fsmwsem_treatsd'^2))/(`secondarysample_fsm_control_n'+`secondarysample_fsm_treat_n
'-2))

local fsmwsem_sdpooled : di %7.2fc `fsmwsem_sdpooled'

local fsmwsem_j =
(exp(lngamma((`secondarysample_fsm_treat_n'+`secondarysample_fsm_control_n'+2)/2)))) /
(((`secondarysample_fsm_treat_n'+`secondarysample_fsm_control_n'+2)/2)^.5 *
exp(lngamma((`secondarysample_fsm_treat_n'+`secondarysample_fsm_control_n'+2-1)/2)))

if `fsmwsem_j' == . {

    local fsmwsem_j = 1-
(3/((4*(`secondarysample_fsm_treat_n'+`secondarysample_fsm_control_n'))-9))

    }

local fsmwsem_g : di %7.2fc `fsmwsem_j'*(`fsmwsem_treatdiff'/`fsmwsem_sdpooled')

local fsmwsem_g_upperci : di %7.2fc
`fsmwsem_j'*(`fsmwsem_treatdiff_upperci'/`fsmwsem_sdpooled')

local fsmwsem_g_lowerci : di %7.2fc
`fsmwsem_j'*(`fsmwsem_treatdiff_lowerci'/`fsmwsem_sdpooled')

* Secondary (Ideation)

quietly summarize ideation_score if secondarysample==1 & FSM==1 & treat==0

local fsmidea_controlsd : di %7.2fc r(sd)

quietly summarize ideation_score if secondarysample==1 & FSM==1 & treat==1

local fsmidea_treatsd : di %7.2fc r(sd)

local fsmidea_sdpooled = sqrt(((`secondarysample_fsm_control_n'-
1)*(`fsmidea_controlsd'^2) + (`secondarysample_fsm_treat_n'-
```

```
1)*(`fsmidea_treatsd'^2))/(`secondarysample_fsm_control_n'+`secondarysample_fsm_treat_n
'-2))

local fsmidea_sdpooled : di %7.2fc `fsmidea_sdpooled'

local fsmidea_j =
(exp(lngamma((`secondarysample_fsm_treat_n'+`secondarysample_fsm_control_n'+2)/2))) /
(((`secondarysample_fsm_treat_n'+`secondarysample_fsm_control_n'+2)/2)^.5 *
exp(lngamma((`secondarysample_fsm_treat_n'+`secondarysample_fsm_control_n'+2-1)/2)))

if `fsmidea_j' == . {

      local fsmidea_j = 1-
(3/((4*(`secondarysample_fsm_treat_n'+`secondarysample_fsm_control_n'))-9))

      }

local fsmidea_g : di %7.2fc `fsmidea_j'*(`fsmidea_treatdiff'/`fsmidea_sdpooled')

local fsmidea_g_upperci : di %7.2fc
`fsmidea_j'*(`fsmidea_treatdiff_upperci'/`fsmidea_sdpooled')

local fsmidea_g_lowerci : di %7.2fc
`fsmidea_j'*(`fsmidea_treatdiff_lowerci'/`fsmidea_sdpooled')

* Missing data analysis - create missing indicator variables

cap drop wam_score_dw_miss

gen wam_score_dw_miss = 0

replace wam_score_dw_miss = 1 if wam_score_dw>=.

logit wam_score_dw_miss i.treat FSM FSMgroup EAL EALgroup group#block, vce(cluster
anonschoolid)

cap drop wsem16_score_miss

gen wsem16_score_miss = 0

replace wsem16_score_miss = 1 if wsem16_score>=.

logit wam_score_dw_miss i.treat FSM FSMgroup EAL EALgroup group#block, vce(cluster
anonschoolid)

* Missing data analysis - only outcome variable missing

regress wam_score_dw i.treat FSM FSMgroup EAL EALgroup group#block, vce(cluster
anonschoolid)

* [IF THERE WERE ANY OTHER STATISTICALLY SIGNIFICANT PREDICTORS OF wam_score_dw_miss
NOT ALREADY HERE WE'D ADD THEM AND REPORT THIS AS A ROBUSTNESS CHECK MODEL]

* Missing data analysis - predictor variables

* [NO MISSING PREDICTOR VARIABLE DATA IN ANALYSIS SAMPLE]

* CREATE TABLES

* Balance table - Randomisation sample

cap file close baltab

file open baltab using output/balance_rsample.csv, write replace

file write baltab '     , , Control, , Intervention' _n

file write baltab 'Pupil-level (categorical), National-level mean, n/N (missing), Count
(%), n/N (missing), Count (%), Standardised difference' _n

file write baltab 'Ever FSM, ${natmean_fsm}, `fsm_controln_rsample' /
`asrandom_schoolcount_control' (`fsm_controlmiss_rsample'), `fsm_controln_in_rsample'
(`fsm_controlpercent_in_rsample'), `

file write baltab ``fsm_treatn_rsample' / `asrandom_schoolcount_treat'
(`fsm_treatmiss_rsample'), `fsm_treatn_in_rsample' (`fsm_treatpercent_in_rsample'),
`fsm_in_stddiff_rsample'`' _n
```

```
file write baltab 'Non-Ever FSM, ${natmean_nonfsm}, `fsm_controln_rsample' /
`asrandom_schoolcount_control' (`fsm_controlmiss_rsample'), `fsm_controln_out_rsample'
(`fsm_controlpercent_out_rsample'), `

file write baltab '`fsm_treatn_rsample' / `asrandom_schoolcount_treat'
(`fsm_treatmiss_rsample'), `fsm_treatn_out_rsample' (`fsm_treatpercent_out_rsample'),
`fsm_out_stddiff_rsample'' _n

file write baltab 'EAL, ${natmean_eal}, `eal_controln_rsample' /
`asrandom_schoolcount_control' (`eal_controlmiss_rsample'), `eal_controln_in_rsample'
(`eal_controlpercent_in_rsample'), `

file write baltab '`eal_treatn_rsample' / `asrandom_schoolcount_treat'
(`eal_treatmiss_rsample'), `eal_treatn_in_rsample' (`eal_treatpercent_in_rsample'),
`eal_in_stddiff_rsample'' _n

file write baltab 'Non-EAL, ${natmean_noneal}, `eal_controln_rsample' /
`asrandom_schoolcount_control' (`eal_controlmiss_rsample'), `eal_controln_out_rsample'
(`eal_controlpercent_out_rsample'), `

file write baltab '`eal_treatn_rsample' / `asrandom_schoolcount_treat'
(`eal_treatmiss_rsample'), `eal_treatn_out_rsample' (`eal_treatpercent_out_rsample'),
`eal_out_stddiff_rsample'' _n

file close baltab

* Balance table – Primary analysis sample

cap file close baltab

file open baltab using output/balance_psample.csv, write replace

file write baltab '    , , Control, , Intervention' _n

file write baltab 'Pupil-level (categorical), National-level mean, n/N (missing), Count
(%), n/N (missing), Count (%), Standardised difference' _n

file write baltab 'Ever FSM, ${natmean_fsm}, `fsm_controln_psample' /
`asrandom_schoolcount_control' (`fsm_controlmiss_psample'), `fsm_controln_in_psample'
(`fsm_controlpercent_in_psample'), `

file write baltab '`fsm_treatn_psample' / `asrandom_schoolcount_treat'
(`fsm_treatmiss_psample'), `fsm_treatn_in_psample' (`fsm_treatpercent_in_psample'),
`fsm_in_stddiff_psample'' _n

file write baltab 'Non-Ever FSM, ${natmean_nonfsm}, `fsm_controln_psample' /
`asrandom_schoolcount_control' (`fsm_controlmiss_psample'), `fsm_controln_out_psample'
(`fsm_controlpercent_out_psample'), `

file write baltab '`fsm_treatn_psample' / `asrandom_schoolcount_treat'
(`fsm_treatmiss_psample'), `fsm_treatn_out_psample' (`fsm_treatpercent_out_psample'),
`fsm_out_stddiff_psample'' _n

file write baltab 'EAL, ${natmean_eal}, `eal_controln_psample' /
`asrandom_schoolcount_control' (`eal_controlmiss_psample'), `eal_controln_in_psample'
(`eal_controlpercent_in_psample'), `

file write baltab '`eal_treatn_psample' / `asrandom_schoolcount_treat'
(`eal_treatmiss_psample'), `eal_treatn_in_psample' (`eal_treatpercent_in_psample'),
`eal_in_stddiff_psample'' _n

file write baltab 'Non-EAL, ${natmean_noneal}, `eal_controln_psample' /
`asrandom_schoolcount_control' (`eal_controlmiss_psample'), `eal_controln_out_psample'
(`eal_controlpercent_out_psample'), `

file write baltab '`eal_treatn_psample' / `asrandom_schoolcount_treat'
(`eal_treatmiss_psample'), `eal_treatn_out_psample' (`eal_treatpercent_out_psample'),
`eal_out_stddiff_psample'' _n

file close baltab

* Outcome measure descriptive statistics

cap file close outcomes
```

```
file open outcomes using output/outcomes.csv, write replace

file write outcomes 'Outcome, Mean, SD, ICC, N' _n

file write outcomes 'WAM Score (ideas scale double weighted), `wam_mean', `wam_sd',
`wam_icc', `wam_n'' _n

file write outcomes 'WSEM-16 Score, `wsem_mean', `wsem_sd', `wsem_icc', `wsem_n'' _n

file write outcomes 'Ideation Score, `idea_mean', `idea_sd', `idea_icc', `idea_n'' _n

file close outcomes

* Primary analysis

cap file close primary

file open primary using output/primary.csv, write replace

file write primary 'Unadjusted means, Full sample , Control group, , Intervention
group, , Effect size calculation, , ' _n

file write primary 'Outcome, n (missing), Mean (95% CI), n (missing), Mean (95% CI), n
(missing), Total n (intervention; control), Hedges' g (95% CI), p-value' _n

file write primary 'WAM Score (ideas scale double weighted), `primarymeasure_n'
(`primarymeasure_miss'), '

file write primary '`wam_controlmean_unadj' (`wam_controlmean_unadj_lowerci';
`wam_controlmean_unadj_upperci'), `primarymeasure_control_n'
(`primarymeasure_control_miss'), '

file write primary '`wam_treatmean_unadj' (`wam_treatmean_unadj_lowerci';
`wam_treatmean_unadj_upperci'), `primarymeasure_treat_n' (`primarymeasure_treat_miss'),
'

file write primary '`primarysample_n' (`primarysample_treat_n';
`primarysample_control_n'), `wam_g' (`wam_g_lowerci'; `wam_g_upperci'),
`wam_treatdiff_p'' _n

file close primary

* Secondary analysis

cap file close secondary

file open secondary using output/secondary.csv, write replace

file write secondary 'Unadjusted means, Full sample , Control group, , Intervention
group, , Effect size, , ' _n

file write secondary 'Outcome, n (missing), Mean (95% CI), n (missing), Mean (95% CI),
n (missing), Total n (intervention; control), Hedges' g (95% CI), p-value' _n

file write secondary 'WSEM Score, `wsem_n' (`wsem_miss'), '

file write secondary '`wsem_controlmean_unadj' (`wsem_controlmean_unadj_lowerci';
`wsem_controlmean_unadj_upperci'), `wsem_control_n' (`wsem_control_miss'), '

file write secondary '`wsem_treatmean_unadj' (`wsem_treatmean_unadj_lowerci';
`wsem_treatmean_unadj_upperci'), `wsem_treat_n' (`wsem_treat_miss'), '

file write secondary '`secondarysample_n' (`secondarysample_treat_n';
`secondarysample_control_n'), `wsem_g' (`wsem_g_lowerci'; `wsem_g_upperci'),
`wsem_treatdiff_p'' _n

file write secondary 'Ideation Score, `idea_n' (`idea_miss'), '

file write secondary '`idea_controlmean_unadj' (`idea_controlmean_unadj_lowerci';
`idea_controlmean_unadj_upperci'), `idea_control_n' (`idea_control_miss'), '

file write secondary '`idea_treatmean_unadj' (`idea_treatmean_unadj_lowerci';
`idea_treatmean_unadj_upperci'), `idea_treat_n' (`idea_treat_miss'), '

file write secondary '`secondarysample_n' (`secondarysample_treat_n';
`secondarysample_control_n'), `idea_g' (`idea_g_lowerci'; `idea_g_upperci'),
`idea_treatdiff_p'' _n
```

```
file close secondary

* Sub-group analyses

cap file close subgroup

file open subgroup using output/subgroup.csv, write replace

file write subgroup 'Model, Hedges' g (95% CI), N, p-value of interaction term, p-value
of treatment variable ' _n

file write subgroup 'WAM FSM sub-group analysis, `fsmwam_g' (`fsmwam_g_lowerci';
`fsmwam_g_upperci'), `primarysample_fsm_n', `fsmwam_interaction_p',
`fsmwam_treatdiff_p'' _n

file write subgroup 'WSEM FSM sub-group analysis, `fsmwsem_g' (`fsmwsem_g_lowerci';
`fsmwsem_g_upperci'), `secondarysample_fsm_n', `fsmwsem_interaction_p',
`fsmwsem_treatdiff_p'' _n

file write subgroup 'Ideation FSM sub-group analysis, `fsmidea_g' (`fsmidea_g_lowerci';
`fsmidea_g_upperci'), `secondarysample_fsm_n', `fsmidea_interaction_p',
`fsmidea_treatdiff_p'' _n

file close subgroup

* Compliance analysis

cap file close subgroup

file open subgroup using output/compliance.csv, write replace

file write subgroup 'Model, Hedges' g (95% CI), N, First stage F test,
Compliance/treatment correlation, p-value of treatment variable ' _n

file write subgroup `'Compliance analysis, `complier_g' '(`complier_g_lowerci',
`complier_g_upperci')', `complier_n', 'F(`complier_df_m', `complier_df_r') =
`complier_F'', `complier_r', `complier_treatdiff_p''' _n

file close subgroup

* Attrition table

cap file close attrition

file open attrition using output/attrition.csv, write replace

file write attrition ' , , Intervention, Control, Total' _n

file write attrition 'Number of pupils, Randomised, `randomisedsample_treat_n',
`randomisedsample_control_n', `randomisedsample_n'' _n

file write attrition '  , Analysed, `primarysample_treat_n', `primarysample_control_n',
`primarysample_n'' _n

file write attrition 'Pupil attrition, Number, `primaryattrition_treat_n',
`primaryattrition_control_n', `primaryattrition_n'' _n

file write attrition '(from randomisation to analysis) , Percentage,
`primaryattrition_treat_p', `primaryattrition_control_p', `primaryattrition_percent''
_n

file close attrition

* Effect size calculation appendix table

cap file close esizecalc

file open esizecalc using output/esizecalc.csv, write replace

file write esizecalc ' , , , Control group, , Intervention group' _n

file write esizecalc 'Outcome, Unadjusted difference in means, Adjusted difference in
means, n (missing), Outcome SD, n (missing), Outcome SD, Pooled SD, Hedges' correction
(J)' _n

file write esizecalc 'WAM Score (ideas double weighted), `wam_treatdiff_unadj',
`wam_treatdiff', '
```

```
file write esizecalc ``primarymeasure_control_n' (`primarymeasure_control_miss'),
`wam_controlsd', `

file write esizecalc ``primarymeasure_treat_n' (`primarymeasure_treat_miss'),
`wam_treatsd', `

file write esizecalc ``wam_sdpooled', `wam_j'' _n

file write esizecalc 'WSEM Score, `wsem_treatdiff_unadj', `wsem_treatdiff', `

file write esizecalc ``wsem_control_n' (`wsem_control_miss'), `wsem_controlsd', `

file write esizecalc ``wsem_treat_n' (`wsem_treat_miss'), `wsem_treatsd', `

file write esizecalc ``wsem_sdpooled', `wsem_j'' _n

file write esizecalc 'Ideation Score, `idea_treatdiff_unadj', `idea_treatdiff', `

file write esizecalc ``idea_control_n' (`idea_control_miss'), `idea_controlsd', `

file write esizecalc ``idea_treat_n' (`idea_treat_miss'), `idea_treatsd', `

file write esizecalc ``idea_sdpooled', `idea_j'' _n

file close esizecalc

* Power Calculation

powercalc, cluster n(`wam_nhmean') g(`wam_schoolcount') icc(`wam_icc')
r2(`wam_r2_pupil') r2_g(`wam_r2_school') treated(`wam_treatprop') blocks(8)
regressors(2)

local wam_mdes : di %7.2fc r(mdes)

powercalc, cluster n(`wsem_nhmean') g(`wsem_schoolcount') icc(`wsem_icc')
r2(`wsem_r2_pupil') r2_g(`wsem_r2_school') treated(`wsem_treatprop') blocks(8)
regressors(2)

local wsem_mdes : di %7.2fc r(mdes)

powercalc, cluster n(`idea_nhmean') g(`idea_schoolcount') icc(`idea_icc')
r2(`idea_r2_pupil') r2_g(`idea_r2_school') treated(`idea_treatprop') blocks(8)
regressors(2)

local idea_mdes : di %7.2fc r(mdes)

* Power Calculation – FSM sub-group

powercalc, cluster n(`wam_fsm_nhmean') g(`wam_schoolcount') icc(`wam_fsm_icc')
r2(`wam_fsm_r2_pupil') r2_g(`wam_fsm_r2_school') treated(`wam_fsm_treatprop') blocks(8)
regressors(2)

local wam_fsm_mdes : di %7.2fc r(mdes)

powercalc, cluster n(`wsem_fsm_nhmean') g(`wsem_schoolcount') icc(`wsem_fsm_icc')
r2(`wsem_fsm_r2_pupil') r2_g(`wsem_fsm_r2_school') treated(`wsem_fsm_treatprop')
blocks(8) regressors(2)

local wsem_fsm_mdes : di %7.2fc r(mdes)

powercalc, cluster n(`idea_fsm_nhmean') g(`idea_schoolcount') icc(`idea_fsm_icc')
r2(`idea_fsm_r2_pupil') r2_g(`idea_fsm_r2_school') treated(`idea_fsm_treatprop')
blocks(8) regressors(2)

local idea_fsm_mdes : di %7.2fc r(mdes)

* Table with inputs for power calculation tables

cap file close powercalc

file open powercalc using output/powercalcinputs.csv, write replace

file write powercalc 'Outcome, MDES, Pupil-level R2, School-level R2, Overall R2, ICC,
Average Cluster Size, Control Schools, Intervention Schools, Total Schools, Control
Pupils, Intervention Pupils, Total Pupils' _n

file write powercalc 'WAM Score, `wam_mdes', `wam_r2_pupil', `wam_r2_school',
`wam_r2_overall', `wam_icc', `wam_nhmean','
```

```
      file write powercalc '`wam_schoolcount_control', `wam_schoolcount_treat',
`wam_schoolcount', `primarysample_control_n', `primarysample_treat_n',
`primarysample_n'' _n

file write powercalc 'WSEM Score, `wsem_mdes', `wsem_r2_pupil', `wsem_r2_school',
`wsem_r2_overall', `wsem_icc', `wsem_nhmean','

      file write powercalc '`wsem_schoolcount_control', `wsem_schoolcount_treat',
`wsem_schoolcount', `secondarysample_control_n', `secondarysample_treat_n',
`secondarysample_n'' _n

file write powercalc 'Ideation Score, `idea_mdes', `idea_r2_pupil', `idea_r2_school',
`idea_r2_overall', `idea_icc', `idea_nhmean', '

      file write powercalc '`idea_schoolcount_control', `idea_schoolcount_treat',
`idea_schoolcount', `secondarysample_control_n', `secondarysample_treat_n',
`secondarysample_n'' _n

file write powercalc 'WAM Score (FSM), `wam_fsm_mdes', `wam_fsm_r2_pupil',
`wam_fsm_r2_school', `wam_fsm_r2_overall', `wam_fsm_icc', `wam_fsm_nhmean','

      file write powercalc '`wam_schoolcount_control', `wam_schoolcount_treat',
`wam_schoolcount', `primarysample_fsm_control_n', `primarysample_fsm_treat_n',
`primarysample_fsm_n'' _n

file close powercalc

* Robustness check including additional controls (time to test, marker fixed effects,
KS1)

rename TestingDate1stVisit TestingDate

regress wam_score_dw i.treat FSM EAL FSMgroup EALgroup group#block if primarysample==1
& TestingDate<., vce(cluster anonschoolid)

estimates store notestdate

estadd scalar es = `wam_j'*(_b[1.treat]/`wam_sdpooled')

regress wam_score_dw i.treat FSM EAL FSMgroup EALgroup group#block TestingDate if
primarysample==1, vce(cluster anonschoolid)

estimates store testdate

estadd scalar es = `wam_j'*(_b[1.treat]/`wam_sdpooled')

recode wam_marker_num .=0

regress wam_score_dw i.treat FSM EAL FSMgroup EALgroup group#block if primarysample==1
& wam_marker_num<., vce(cluster anonschoolid)

estimates store nomarker

estadd scalar es = `wam_j'*(_b[1.treat]/`wam_sdpooled')

regress wam_score_dw i.treat FSM EAL FSMgroup EALgroup group#block i.wam_marker_num if
primarysample==1, vce(cluster anonschoolid)

estimates store marker

estadd scalar es = `wam_j'*(_b[1.treat]/`wam_sdpooled')

recode wam_marker_num 0=.

mi set flong

mi register imputed tks1ave_round

mi impute regress tks1ave_round FSM EAL FSMgroup EALgroup group#block if
primarysample==1, add(20) rseed(11849)

mi estimate, post: regress wam_score_dw i.treat FSM EAL FSMgroup EALgroup group#block
if primarysample==1 & tks1ave_round<., vce(cluster anonschoolid)

estimates store noks1

estadd scalar es = `wam_j'*(_b[1.treat]/`wam_sdpooled')
```

```
mi estimate, post: regress wam_score_dw i.treat FSM EAL FSMgroup EALgroup group#block
tks1ave_round if primarysample==1, vce(cluster anonschoolid)

estimates store ks1

estadd scalar es = `wam_j'*(_b[1.treat]/`wam_sdpooled')

mi extract 0, clear

estimates restore primary

estimates store primary

estadd scalar es = `wam_g'

estout primary notestdate testdate nomarker marker noks1 ks1 using
output/robustness.csv, replace ///
      mlabels('Primary' 'No Test Date' 'Test Date' 'No Marker FE' 'Marker FE' 'No KS1'
'KS1', nonumbers nodepvars) ///
      keep(1.treat) varlabels(1.treat 'Treatment') ///
      indicate('Test Date = TestingDate' 'Marker FE = 4.wam_marker_num' 'KS1 =
tks1ave_round') ///
      cells(b(fmt(3) star) p(fmt(3) par)) delimiter(', ') ///
      stats(es N, fmt(2 0) labels('Effect Size' 'N'))

* Randomisation inference as a robustness check

cap drop groupblock

egen groupblock = concat(group block)

encode groupblock, gen(groupblock_enc)

drop groupblock

rename groupblock_enc groupblock

ritest treat _b[1.treat], r(2000) strata(groupblock) cluster(anonschoolid) seed(987234)
: ///
    regress wam_score_dw i.treat FSM EAL FSMgroup EALgroup group#block, vce(cluster
anonschoolid)

local wam_ri_p : di %7.2fc r(p)[1,1]

ritest treat _b[1.treat], r(2000) strata(groupblock) cluster(anonschoolid) seed(987234)
: ///
    regress wsem16_score i.treat FSM EAL FSMgroup EALgroup group#block, vce(cluster
anonschoolid)

local wsem_ri_p : di %7.2fc r(p)[1,1]

ritest treat _b[1.treat], r(2000) strata(groupblock) cluster(anonschoolid) seed(987234)
: ///
    regress ideation_score i.treat FSM EAL FSMgroup EALgroup group#block, vce(cluster
anonschoolid)

local idea_ri_p : di %7.2fc r(p)[1,1]

* Table to compare p values from randomisation inference and sampling inference

cap file close ricomp

file open ricomp using output/ricomp.csv, write replace

file write ricomp 'Outcome, Effect size, Sampling Inference p-value, Randomisation
Inference p-value' _n

file write ricomp 'WAM Score (ideas scale double weighted), `wam_g', `wam_treatdiff_p',
`wam_ri_p'' _n

file write ricomp 'WSEM Score, `wsem_g', `wsem_treatdiff_p', `wsem_ri_p'' _n
```

```
file write ricomp 'Ideation Score, `idea_g', `idea_treatdiff_p', `idea_ri_p'' _n

file close ricomp

* Table with number of treatment and control schools in each block

cap file close strata

file open strata using output/strata.csv, write replace

file write strata 'Randomisation Batch, Stratum, Treatment (N), Control (N)' _n

file write strata '1, Low EAL/Low FSM, `block11treat_schoolcount_1',
`block11control_schoolcount_1'' _n

file write strata '1, Low EAL/High FSM, `block12treat_schoolcount_1',
`block12control_schoolcount_1'' _n

file write strata '1, High EAL/Low FSM, `block21treat_schoolcount_1',
`block21control_schoolcount_1'' _n

file write strata '1, High EAL/High FSM, `block22treat_schoolcount_1',
`block22control_schoolcount_1'' _n

file write strata '2, Low EAL/Low FSM, `block11treat_schoolcount_2',
`block11control_schoolcount_2'' _n

file write strata '2, Low EAL/High FSM, `block12treat_schoolcount_2',
`block12control_schoolcount_2'' _n

file write strata '2, High EAL/Low FSM, `block21treat_schoolcount_2',
`block21control_schoolcount_2'' _n

file write strata '2, High EAL/High FSM, `block22treat_schoolcount_2',
`block22control_schoolcount_2'' _n

file close strata
```

Appendix G: Teacher interview guide

Young Journalist Academy: Interviews with classroom teacher

The interviews should last around 30 minutes. The timings given for each section are a guide – you may spend longer or shorter on each section. Lead questions are presented in bold, with potential follow-up questions presented in a non-bold typeface. As the interviews are semi-structured, not all questions need to be asked and they do not need to be asked in order. The interviewer should be responsive to what the interviewee, following the direction of the conversation and following-up with additional questions as needed.

Main objective	Purpose of section	Guide timings
1. Introduction	Explains the purpose and 'ground rules' of the interview.	3 mins
2. Background context	Allows the participant an opportunity to settle into the interview, as well as providing some background to the school and the context in which the Young Journalist Academy programme is being delivered.	5 mins
3. Delivery experience	This section will focus on understanding the perceived quality of the intervention, as well as experiences of the programme's delivery, including barriers and facilitators to delivery.	5 mins
4. Pupil engagement and programme mechanisms	To explore the teacher's perception of pupils' engagement in the sessions and the positive and negative impact of the programme, together with the mechanisms that brought about any impact identified.	5 mins
5. Support and dissemination	To understand any support that teachers have received to deliver the programme, including from the school's SLT, and how they have disseminated the programme within the school.	5 mins
6. Close	Thank you and close	2 mins
Observation of YJA session		
7. Reflection following observation of YJA session	Where possible, this will be an opportunity to follow-up on any areas of interest arising from the observation.	5 mins

1. Introduction	3 mins

Introduction: ● Introduce yourself ● Introduce BIT and IOE – explain that we are independently evaluating the Young Journalist Academy programme, which is one of five programmes that are part of the Cultural Learning programme that is jointly funded by the Education Endowment Foundation and Royal Society of Arts. **Aims of this interview:** We are here to learn more about how the Young Journalist Academy programme has worked in your class. We're interested in what involvement you have had with the programme and what has helped the programme to work, and what the challenges have been. We'd also like to understand any impact the programme has had on your school, particularly pupils in your class. **This interview:** ● Should take no more than 30 minutes ● Stress that you want to understand the intervention from their point of view. No answers are right or wrong – and we are not here to judge the decisions made or views held by the interviewee. **Anonymity and privacy:** ● All information gathered will be in strict confidence, unless there are concerns about safeguarding. When we write up the research we will ensure that no one is identifiable from any reporting. ● Explain that if at any point they feel uncomfortable or prefer not to answer a specific question they can just say so. ● Explain that it is their choice whether they take part in the interview and they can end the interview at any point, without giving a reason. **Recording:** 1. Explain that recording enables us to have an accurate record of what was said, which can be typed up for analysis alongside other interviews. We may also use quotes from this interview, but these will be included in a way that means no individual or school is identifiable. 2. Check if they have any questions about the interview. If they are happy to go ahead, obtain verbal permission to digitally record and take notes (written permission should already have been obtained). 3. Once you have consent, start the voice recorder. 4. State interview number/participant ID	Orientates respondent and gets them prepared to take part in the discussion. Outlines the 'rules' of the interview.
2. Background context	5 mins

How many years have you been teaching? **How long have you been working at this school?** **Could you tell me a little bit about your role at the school?** ● What would you say are the school's main strengths and challenges? **How did you become involved with the Young Journalist Academy project?** **What were your thoughts about the Young Journalist Academy programme when you first heard about it?** ● How did you feel about your class taking part in the Young Journalist Academy sessions **When did you begin your Young Journalist Academy sessions?** ● What did the initial set up process involve? ● How did you find the initial set up process? **How is writing typically taught to pupils in Year 5 at your school?** ● How does the Young Journalist Academy compared to this? ● Have students previously engaged in journalistic writing activities? If yes, what have they done?	Allows the participant an opportunity to settle into the interview, as well as providing some background to the school and the context in which the Young Journalist Academy programme is being delivered.
3. Delivery experience	5 mins

How has the Young Journalist Academy been organised and timetabled in your curriculum?
 	
● When in the school year did you timetable the sessions? Why was this?
 	
● What did you think about the ordering/sequencing of the session content?
 	
● What do you think about the typical length of sessions?
 	
● Where do the sessions typically take place? How have you found using this space for the sessions?
 	
● Were there any particular resources or technology you had to provide for the sessions?
 	
How have you found the Young Journalist Academy sessions?
 	
● What do you like about the Young Journalist Academy programme?
 	
   • [probe for thoughts on content of sessions, quality of delivery (e.g. by external provider), student engagement with material]
 	
● What have the main challenges been?
 	
   • [probe for: timetabling, working with external provider, student engagement, facilities, resources, behaviour management]
 	
● What do you think could have helped to overcome these?
 	
● Can you tell me any recommendations you have for how the programme could be improved?
 	
How have you found working with the Young Journalist Academy Mentor?
 	
● How have they been with the pupils?
 	
● To what extent have you been able to share your knowledge with one another?
 	
● Did you face any challenges working with the YJA mentor?
 	
● How did you overcome those challenges?
 	
● What is your typical role within the sessions?
 	
● What has your role been in the editing process? How have you found this?
 	
Can you tell me a bit about the mid-term review, when a member of YJA came in to the school to see how the programme was running?
 	
● Can you tell me what you discussed at the review?
 	
● What actions, if any, came out of the review?
 	
● Can you tell me about anything that was useful about the review?
 	
● Can you tell me anything you would change about the review?
	This section will focus on understanding the perceived quality of the intervention, as well as experiences of the programme's delivery, including barriers and facilitators to delivery.

4. Pupil Engagement and Programme Mechanisms	5 mins

How have the children in your class responded to the Young Journalist Academy sessions? ● Could you describe the pupils who have got the most out of the sessions? Why do you think this is? ● Could you describe any children who have struggled more with the sessions? Why do you think this is? **As I understand, a small group of students participating in the YJA programme are selected to be part of a 'core' editorial team. What did being part of the 'core' team involve for these students?** ● How were students selected for this role? ● Could you describe any additional impacts for these students from being part of the 'core' team? ● What was the relationship between the 'core' team and the rest of the class? ● How did this affect engagement with the programme? **More generally, what effects have the Young Journalist Academy sessions had for your pupils?** ● Can you describe any other changes you've noticed in your pupils? ● [If not mentioned, probe for class engagement with literacy, writing ability, self efficacy (reading, writing, listening), motivation, ability to critically analyse information, self-regulation] ● What is it about programme that you think has helped them to develop their (e.g. classroom motivation)? ● Can you describe any negative consequences of the programme for pupils? **Can you describe any ways that taking part in the Young Journalist Academy has affected you?** ● [probe for: classroom practice, workload, curriculum links] ● What was it about the programme that you think led to these changes? ● Can you describe any skills or techniques explored in the Young Journalist Academy sessions that you used with your class? **Could you tell me about some of the content students have produced as part of the Young Journalist Academy programme?** ● How have you disseminated their work? In school? More widely?	To explore the the teacher's perception of pupils' engagement in the sessions and the positive and negative impact of the programme, together with the mechanisms that brought about any impact identified..
5. Support and dissemination	5 mins

How have the SLT responded to the Young Journalist Academy programme? ● Can you describe any support you've received from SLT to deliver the programme? ● Can you describe any resources they've made available to deliver the programme? **Can you describe any impacts participating in the YJA programme has had on the wider school?** ● How have you shared the content of the YJA programme with other teachers students within the school? What has the impact of this been? ● How have you shared the content of the YJA programme with other students within the school? What has the impact of this been?	To understand teacher's experience of the training they received to deliver the programme, and the support they have received to deliver the programme, including from the school's SLT.
6. Close	2 mins
Overall, would you recommend Young Journalist Academy to other schools? ● Why/Why not? **Was there anything else that you were hoping to discuss that we haven't yet had a chance to talk about?** Thank the interviewee for their time and reassure them of the confidentiality of their responses, as explained at the beginning of the interview.	Thank you and close.

7. Reflection following observation of YJA session	5 mins
Below are some exemplar questions, however, the interviewer may want to add some additional questions based on their observations. Probe for their perception of student engagement; implementer support; and mechanisms. **How did you find that session went?** ● What do you think went well? ● What do you think the challenges were? ● Which students engaged particularly well? How/why? ● Which students engaged less well? How/why? If you could go back and do that session again, what might you do differently? **How does that session compare to a typical Young Journalist Academy session?** Probe specific areas of interest from your observation/ in relation to the review questions/ thinking about how this school is running Young Journalist Academy in comparison to other schools. **Could you describe any ways that the sessions have changed over time?**	Opportunity to follow-up on any areas of interest arising from the observation.

Appendix H: Senior leadership team (SLT) interview guide

Young Journalist Academy: Interviews with senior leadership team

The interviews should last around 30 minutes. The timings given for each section are a guide – you may spend longer or shorter on each section. Lead questions are presented in bold, with potential follow-up questions presented in a non-bold typeface. As the interviews are semi-structured, not all questions need to be asked and they do not need to be asked in order. The interviewer should be responsive to what the interviewee, following the direction of the conversation and following-up with additional questions as needed.

Main objective	Purpose of section	Guide timings
1. Introduction	Explains the purpose and 'ground rules' of the interview.	3 mins
2. Background context	Allows the participant an opportunity to settle into the interview, as well as providing some background to the school, so that we understand more about the context in which the YJA programme is being delivered.	4 mins
3. Engagement	To understand the extent to which the school, and particularly the SLT, has engaged with and supported the programme.	7 mins
4. Delivery experience	This section will focus on understanding the perceived quality of the intervention, as well as the experience of the programme's delivery, including barriers and facilitators to delivery. We will also try and disentangle what is 'business as normal' within the school in terms of writing, and whether this has changed as a result of the YJA programme.	7 mins
5. Mechanisms of change	To explore the SLT member's perception of the positive and negative impact of the programme, particularly for pupils, and the mechanisms that brought about any change identified.	7 mins
6. Close	Thank you and close	2 mins

1. Introduction	3 mins

Introduction: ● Introduce yourself ● Introduce BIT and IOE – explain that we are independently evaluating the First Thing Music programme, which is one of five programmes that are part of the Cultural Learning programme that is jointly funded by the Education Endowment Foundation and Royal Society of Arts. **Aims of this interview:** We are here to learn more about how the YJA programme has worked in your school. We're interested in what involvement you have had with the programme, what has helped the programme to work, and what the challenges have been. We'd also like to understand any impact the programme has had on your school, particularly pupils in (the) YJA class(es). **This interview:** ● Should take no more than 20 minutes ● Stress that you want to understand the intervention from their point of view. No answers are right or wrong – and we are not here to judge the decisions made or views held by the interviewee. **Anonymity and privacy:** ● All information gathered will be in strict confidence, unless there are concerns about safeguarding. When we write up the research we will ensure that no one is identifiable from any reporting. ● Explain that if at any point they feel uncomfortable or prefer not to answer a specific question they can just say so. Explain that it is their choice whether they take part in the interview and they can end the interview at any point, without giving a reason **Recording:** 1. Explain that recording enables us to have an accurate record of what was said, which can be typed up for analysis alongside other interviews. We may also use quotes from this interview, but these will be included in a way that means no individual or school is identifiable. 2. Check if they have any questions about the interview. If they are happy to go ahead, obtain verbal permission to digitally record and take notes (written permission should already have been obtained). 3. Once you have consent, start the voice recorder. 4. State interview number/participant ID	Orientates respondent and gets them prepared to take part in the discussion. Outlines the 'rules' of the interview.
2. Background context	4 mins

How many years have you been teaching?
 	
 How long have you been working at the school?
 	
 Could you tell me about your role at the school?
 	
 Could you tell me about what the school is like?
 	
     ● What would you say are the schools' main strengths?
 	
     ● What would you say are some of the school's biggest challenges?
 	
 How did your school become involved with the Young Journalist Academy programme?
 	
     ● When did you first hear about the Young Journalist Academy programme?
 	
     ● Why did your school decide to get involved in the programme?
 	
     ● What initial expectations did you have for the programme?
	Allows the participant to settle into the interview, as well as providing some background to the school, so that we understand more about the context in which the Young Journalist Academy programme is being delivered.

3. Engagement	7 mins

Can you describe what your involvement with the Young Journalist Academy programme has been?
 	
     ● Can you describe any support that you or other members of SLT have provided to the programme?
 	
     ● Can you describe any resources that you've made available to enable the programme to run?
 	
 Have you sat in on any Young Journalist Academy sessions?
 	
 If yes:
 	
     ● Can you describe any sessions you've sat in on?
 	
     ● What did you think of the session(s)?
 	
 [probe for further exploration of strength and suggested areas for improvement]
 	
  If no:
 	
     ● Were there any particular reasons that you didn't sit in on a session?
 	
     ● What do you understand about what the class have been doing in their Young Journalist Academy sessions?
	To understand the extent to which the school, and particularly the SLT, has engaged with and supported the programme

4. Delivery experience	7 mins

How is writing typically taught to pupils in Year 5 at your school? ● What writing interventions would pupils in [intervention year] typically receive? ● Have students previously engaged in journalistic writing activities? ● How does the Young Journalist Academy compared to this? **Thinking specifically about the delivery of the programme, how have you found having the Young Journalist Academy programme running in your school?** ● What has worked well about the delivery of the programme? [probe for thoughts on content of sessions, quality of delivery (e.g. by external provider), student engagement with material] ● What have been the main challenges to delivering the programme? [probe for: timetabling issues, working with external partner, behaviour management, student engagement, facilities, resources] ● What have you done to try and overcome these challenges? ● How has the school found accommodating the Young Journalist Academy sessions into the school timetable? [e.g 8 sessions across the year – probe for which content it replaces in curriculum] ● Knowing what you know now, would you approach anything about the programme implementation differently	This section will focus on understanding the perceived quality of the intervention, as well as the experience of the programme's delivery, including barriers and facilitators to delivery. We will also try and disentangle what is 'business as normal' within the school in terms of writing, and whether this has changed as a result of the Young Journalist Academy programme.

5. **Mechanisms of change**	7 mins
How have pupils in [intervention class] found the Young Journalist Academy programme? ● To what extent have pupils engaged with the sessions? ● What effect do you think taking part in the Young Journalist Academy programme has had for pupils? [probe for logic model outcomes: ability to critically analyse writing, self-regulation, increased motivation, self regulation, self-efficacy) ● What was it about the programme that you think led to [change described]? ● Can you describe any pupils that you think the programme is particularly helpful for? [probe for core team vs. rest of class] ● Can you describe any pupils that you think the programme is less suited to? **What effect, if any, do you think that the programme has had for the teacher of [intervention class]?** ● Can you think of any other impacts that the programme has had for the [intervention class] teacher? **Can you describe any effect of the programme more broadly within the school?** ● Can you describe any effect of the programme on other pupils within the school? ● Can you describe any effect of the programme on other teachers within the school?	To explore the SLT member's perception of the positive and negative impact of the programme, particularly for pupils, and the mechanisms that brought about any change identified.

● Has [intervention class] doing Young Journalist Academy sessions affected writing provision and students' engagement with literacy within the school more generally? If yes, how?	
6. Close	2 mins
Overall, would you recommend Young Journalist Academy to other schools? ● Why/Why not? **Was there anything else that you were hoping to discuss that we haven't yet had a chance to talk about?** Thank the interviewee for their time and reassure them of the confidentiality of their responses, as explained at the beginning of the interview.	Thank you and close

Appendix I: Young Journalist Academy mentor interview guide

Young Journalist Academy: Interviews with Young Journalist Academy mentor

The interviews should last around 30 minutes. The timings given for each section are a guide – you may spend longer or shorter on each section. Lead questions are presented in bold, with potential follow-up questions presented in a non-bold typeface. As the interviews are semi-structured, not all questions need to be asked and they do not need to be asked in order. The interviewer should be responsive to what the interviewee, following the direction of the conversation and following-up with additional questions as needed.

Main objective	Purpose of section	Guide timings
1. Introductions	Explains the purpose and ground rules for the interview.	3 mins
2. Background context	Allows the participant an opportunity to settle into the interview, as well as providing some background to the school and the context in which the YJA programme is being delivered.	5 mins
3. Delivery experience	This section will focus on understanding the perceived quality of the intervention, as well as experiences of the programme's delivery, including barriers and facilitators to delivery.	5 mins
4. Pupil engagement and programme mechanism	To explore the drama practitioner's perception of pupils' engagement in the sessions and the positive and negative impact of the programme, together with the mechanisms that brought about any impact identified.	5 mins
5. Training and support	To understand drama practitioner's experience of the training they received to deliver the programme, and the support they have received to deliver the programme, including from the school's SLT and class teachers.	5 mins
6. Close	Thank you and close.	2 mins
Observation of YJA session		
7. Reflection following observation of YJA session	Where possible, this will be an opportunity to follow-up on any areas of interest arising from the observation.	5 mins

Topic guide

1. Introductions	5 mins

Introduction: ● Introduce yourself ● Introduce BIT and IOE – explain that we are independently evaluating the First Thing Music programme, which is one of five programmes that are part of the Cultural Learning programme that is jointly funded by the Education Endowment Foundation and Royal Society of Arts. **Aims of this interview:** We are here to learn more about how the YJA programme has worked in your class. We're interested in what involvement you have had with the programme and what has helped the programme to work, and what the challenges have been. We'd also like to understand any impact the programme has had on your school, particularly pupils in your class. **This interview:** ● Should take no more than 30 minutes Stress that you want to understand the intervention from their point of view. No answers are right or wrong – and we are not here to judge the decisions made or views held by the interviewee. **Anonymity and privacy:** ● All information gathered will be in strict confidence, unless there are concerns about safeguarding. When we write up the research we will ensure that no one is identifiable from any reporting ● Explain that if at any point they feel uncomfortable or prefer not to answer a specific question they can just say so. ● Explain that it is their choice whether they take part in the interview and they can end the interview at any point, without giving a reason. **Recording:** 1. Explain that recording enables us to have an accurate record of what was said, which can be typed up for analysis alongside other interviews. We may also use quotes from this interview, but these will be included in a way that means no individual or school is identifiable. 2. Check if they have any questions about the interview. If they are happy to go ahead, obtain verbal permission to digitally record and take notes (written permission should already have been obtained). 3. Once you have consent, start the voice recorder.	Orientates respondent and gets them prepared to take part in the discussion. Outlines the 'rules' of the interview.
2. Background context	5 mins

How did you become involved with the Young Journalist Academy project? **What were your thoughts about the Young Journalist Programme programme when you first heard about it?** ● How did you feel about leading the Young Journalist Academy project at [case study school]? ● What were you looking forward to about your taking part? ● What challenges did you envisage? **Have you previously been involved in any similar programmes?** ● How does Young Journalist Academy compare to these programmes? **How would you describe this school?** ● Are you aware of any particular challenges faced by the school?	Allows the participant an opportunity to settle into the interview, as well as providing some background to the school and the context in which the YJA programme was delivered.
3. Delivery Experience	5 mins
How have you found delivering the Young Journalist Academy sessions? ● What has the Young Journalist Academy programme consisted of? ● What is the *role and purpose* of the YJA mentor within the Young Journalist Academy programme? ● What has *worked well* about delivering the Young Journalist Academy programme at this school? ● What have the main *challenges* been? ● What have you done to overcome these challenges? ● How suitable have the schools *facilities and resources* for the Young Journalist Academy sessions? [probe for physical facilities (e.g. classroom space) and also resources (e.g. ICT equipment, stationary)] ● Knowing what you know now, would you approach anything about the programme implementation within [case study school] differently? **How did you find working with the classroom teacher?** ● What was their typical role within the session? ● Did you face any challenges working with them? ● How did you overcome these challenges? ● How did you support them to engage with Young Journalist Academy? ● To what extent have you been able to share your knowledge with one another?	This section will focus on understanding the perceived quality of the intervention, as well as experiences of the programme's delivery, including barriers and facilitators to delivery.
3. Pupil Engagement and Programme Mechanisms	10 mins

How have children in your sessions found the Young Journalist Academy sessions? ● Could you describe the pupils who have got the most out of the sessions? Why do you think this is? ● Could you describe any children who have struggled more with the sessions? Why do you think this is? **As I understand, a small group of students participating in the YJA programme are selected to be part of a 'core' editorial team. What did being part of the 'core' team involve for these students?** ● How were students selected for this role? ● Could you describe any additional impacts for these students from being part of the 'core' team? ● What was the relationship between the 'core' team and the rest of the class? ● How did this affect engagement with the programme? **More generally, what effects have the Young Journalist Academy sessions had for pupils at [case study school]?** ● Can you describe any other changes you've noticed in your pupils? ● [If not mentioned, probe for class engagement with literacy, writing ability, self efficacy (reading, writing, listening), motivation, ability to critically analyse information, self-regulation] ● What is it about programme that you think has helped them to develop their (e.g. classroom motivation)? ● Can you describe any negative consequences of the programme for pupils? **To what extent did you adapt your sessions for students of different abilities or with different needs?** ● Did you use any particular physical resources as part of your Young Journalist Academy sessions? **How did your experiences at [case study school] compare to the other schools you have worked with?** ● How similar/different were the schools you worked with? ● How have you adapted the Young Journalist Academy programme to support the needs of different schools? ● Are there any learnings from other schools you worked with that you would like to share with us?	To explore the drama practitioner's perception of pupils' engagement in the sessions and the positive and negative impact of the programme, together with the mechanisms that brought about any impact identified.
5. Training and Support	5 mins

(Less vital to ask about**) Can you tell me about the training sessions that you attended as part of the Young Journalist Academy programme?** ● What did you cover in the training sessions? ● What worked well about the sessions? ● What did you find challenging about the sessions? ● To what extent did the sessions support you to deliver the Young Journalist Academy programme? ● If you could change anything about the training, what would it be? **How have the SLT responded to the Young Journalist Academy programme at [case study school]?** ● Can you describe any support you've received from SLT to deliver the programme? ● Can you describe any resources they've made available to deliver the programme?	To understand drama practitioner's experience of the training they received to deliver the programme, and the support they have received to deliver the programme, including from the school's SLT and class teachers.
6. Close	2 mins
If you had a magic wand, what one thing would you change about the Young Journalist Academy and why? **Was there anything else that you were hoping to discuss that we haven't yet had a chance to talk about?** Thank the interviewee for their time and reassure them of the confidentiality of their responses, as explained at the beginning of the interview.	Thank you and close

7. Reflection following observation of YJA session	5 mins
Below are some exemplar questions, however, the interviewer may want to add some additional questions based on their observations. Probe for their perception of student engagement; implementer support; and mechanisms. **How did you find that session went?** ● What do you think went well? ● What do you think the challenges were? ● Which students engaged particularly well? How/why? ● Which students engaged less well? How/why? If you could go back and do that session again, what might you do differently? **How does that session compare to a typical Young Journalist Academy session?** Probe specific areas of interest from your observation/ in relation to the review questions/ thinking about how this school is running Young Journalist Academy in comparison to other schools. **Could you describe any ways that the sessions have changed over time?**	Opportunity to follow-up on any areas of interest arising from the observation.

Appendix J: Guide for informal discussions with pupils

Young Journalist Academy: Discussions with pupils

Short discussions, lasting around five minutes will be conducted with pupils happy to speak to a researcher. These will happen during lesson time in the classroom setting. The researcher(s) will take fieldnotes and therefore the conversations will not be audio recorded.

Aims of discussions with pupils

To better understand pupils 'responses to and engagement with the Young Journalist Academy programme.

Before conducting discussions

● The researcher should find out from the teacher in advance which pupils 'parents have consented for them to take part in the research, and whether there are pupils who have particular communication needs, where the researchers will need to adapt their approach.

● The teacher should explain to the class: who the researcher is and what they are going to be doing; the purpose of the research; and reassure pupils that they do not have to talk to the researcher if they would prefer not to.

● Before speaking to a pupil, ask whether they are okay to speak to you and ensure that their body language indicates that they are happy to talk to you.

Discussion guide questions

How are your finding your YJA session today?

● What do you like about it?

● What do you not like about it?

Is there anything you would like to do differently/do more of/do less of in your Young Journalist Academy session?

Can you tell me about some of the things you've learnt in your Young Journalist Academy sessions?

Could you tell me about some of the articles you've written or helped to write as part of the Young Journalist Academy programme?

● How did you find writing these articles?

How did you decide what to write about?

What do you think of the Young Journalist Academy mentor who comes in to lead the Young Journalist Academy sessions?

Is there anything else you do in school that's like the Young Journalist Academy or is the Young Journalist Academy very different from everything else?

·   Like: literacy lessons/ school newspaper – could you tell me about them?

·   Different: real life element/role of teacher/external speaker/literacy lessons – could you tell me about them?

Is there anything else you want to tell me about the Young Journalist Academy programme?

Thank you for talking to me today.

Appendix K: Observation proforma

Young Journalist Academy: Observation proforma

This proforma may be completed during the observation, or more extensive notes may be made during the observation, i.e. in a notepad, and then summarised on this form.

Date and time: Name of school and class: Number of pupils present: Name of observer: Other adults present: Notes		
Introduction	What is covered in the introduction to the session, e.g. are the aims and expectations covered?	
Activities	What activities are carried out as part of the session?	
	What kind of journalistic outputs do students produce? TWE do they have ownership of their work (are they given choice at different stages of the writing process? Are they proud of their work?) Do they receive feedback? How do students respond to feedback? Are the children given time to write freely? Is there any discussion about language choices and how they affect writing? If yes, what is discussed?	
Teacher/Mentor Behaviour	What is the role of the classroom teacher within the session? How are tasks shared? How is knowledge shared?	
Pupil Engagement /Behaviour	To what extent are children engaged in the session? e.g. are they focused on writing, or are they distracted and talking to their peers about unrelated things?	
	What is the behaviour of the pupils like?, e.g. mainly on-task; mainly disruptive; individual children not on-task but the majority engaged? What strategies are used to manage pupils behaviour?	
	What strategies are used to account for the needs and abilities of different pupils? What difficulties do children	

	who struggle seem to have? How are they supported?	
	What is the role of the 'core' team within the session? How are tasks allocated between the 'core' team and the rest of the group? What is the relationship between the 'core' team and the rest of the class?	
Materials/ Resources	What resources/technology are used as part of the session?How do these impact students' engagement? What journalistic materials and information are students exposed to as part of the session? How do students respond to these?	
Any other observations?		

Appendix L: Raw treatment condition survey data

Please note: some percentages may add up to slightly above or below 100% due to rounding.

1.1 What is your role at the school?	
Response	**Number (%)**
Class teacher	33 (94%)
Other members of the SLT	3 (9%)
Teaching assistant	1 (3%)
Head Teacher	0 (0%)
Deputy head teacher	0 (0%)
Other	0 (0%)

N.B. Total percentage is greater than 100% because respondents could select more than one role.

2.1 What activities did your class take part in during the eight Young Journalist Academy teaching days (when the YJA mentor came into your school)?	
Response	**Number (%)**
Article writing	35 (100%)
Radio production	35 (100%)
TV production	32 (91%)
Other (please specify): - TV production via iPads	1 (3%)
Don't know	0 (0%)

N.B. Total percentage is greater than 100% because respondents could select more than one answer

2.2 Of the eight Young Journalist Academy teaching days, approximately how many days did you spend on: Article writing

Response	Number (%)
4 days	16 (46%)
5 days	6 (17%)
3 days	5 (14%)
2 days	4 (11%)
6 days	3 (9%)
7 days	1 (3%)

2.3 Of the eight Young Journalist Academy teaching days, approximately how many days did you spend on: Radio Production

Response	Number (%)
2 days	27 (77%)
1 day	4 (11%)
3 days	4 (11%)

2.4 Of the eight Young Journalist Academy teaching days, approximately how many days did you spend on: TV production

Response	Number (%)
2 days	21 (60%)
1 day	11 (31%)
0 days	2 (6%)
3 days	1 (3%)

3.1 To what extent do you agree or disagree with the following statement: I found the Young Journalist Academy mentor easy to work with

Response	Number (%)
Strongly agree	19 (54%)
Agree	8 (23%)
Neither agree nor disagree	6 (17%)
Disagree	1 (6%)
Strongly disagree	0 (0%)
Don't know	0 (0%)

3.2 To what extent do you agree or disagree with the following statement: The Young Journalist Academy mentor supported me to deliver lessons/activities outside the eight Young Journalist Academy teaching days

Response	Number (%)
Strongly agree	5 (14%)
Agree	9 (26%)
Neither agree nor disagree	14 (40%)
Disagree	6 (17%)
Strongly disagree	0 (0%)
Don't know	1 (3%)

3.3 How would you rate the YJA mentor's delivery of the programme to your class?

Response	Number (%)
Very good	18 (51%)

Good	6 (17%)
OK	9 (26%)
Poor	1 (3%)
Very poor	1 (3%)
Don't know	0 (0%)

4.1 Outside the eight-day Young Journalist Academy teaching days, how often did you involve your class in Young Journalist Academy related activities?

Response	Number (%)
Once a week	10 (29%)
Once a month	8 (23%)
Less than monthly	7 (20%)
Never	3 (9%)
Several times a week	1 (3%)
Other (please specify):	
- Once a fortnight	1 (3%)
- Several times a month	1 (3%)
- Varied depending on how much work we had to complete	1 (3%)
- About weekly	1 (3%)
- On two occasions	1 (3%)
- Follow-up lessons after each YJA session	1 (3%)

4.2 Which of the following activities have you done with your class, outside the eight Young Journalist Academy teaching days?

Response	Number (%) answered 'Yes'
Writing newspaper articles	22 (63%)
Watching the news/news clips	21 (60%)

Reading newspaper articles/news online	20 (57%)
Incorporating the editorial feedback process into lessons/class activities	9 (26%)
Listening to radio programmes/podcasts	8 (23%)
Using the Young Journalist Academy website in lessons/class activities	7 (20%)
Critical analysis of media outputs	6 (17%)
Producing news/video clips	5 (14%)
Producing radio programmes/podcasts	1 (3%)

N.B. Total percentage is greater than 100% because respondents could select more than one answer

4.3 Which of the following factors affected (either positively or negatively) your decision to do Young Journalism related activities outside the eight Young Journalist Academy teaching days? Please rate each of the factors from 1-5 where 1=not influential at all, and 5=highly influential.

Factor: Children's engagement with Young Journalist Academy activities

Response	Number (%)
5 Highly influential	9 (26%)
4	12 (34%)
3	12 (34%)
2	1 (3%)
1 Not influential at all	1 (3%)

4.4 Which of the following factors affected (either positively or negatively) your decision to do Young Journalism related activities outside the eight Young Journalist Academy teaching days? Please rate each of the factors from 1-5 where 1=not influential at all, and 5=highly influential.

Factor: Time in the curriculum

Response	Number (%)
5 Highly influential	15 (43%)

4	5 (14%)
3	8 (23%)
2	1 (3%)
1 Not influential at all	6 (17%)

4.5 Which of the following factors affected (either positively or negatively) your decision to do Young Journalism related activities outside the eight Young Journalist Academy teaching days? Please rate each of the factors from 1-5 where 1=not influential at all, and 5=highly influential.

Factor: Fit with other learning objectives for my class

Response	Number (%)
5 Highly influential	11 (31%)
4	12 (34%)
3	7 (20%)
2	4 (11%)
1 Not influential at all	1 (3%)

4.6 Which of the following factors affected (either positively or negatively) your decision to do Young Journalism related activities outside the eight Young Journalist Academy teaching days? Please rate each of the factors from 1-5 where 1=not influential at all, and 5=highly influential.

Factor: Perception of Young Journalist Academy impact on children's learning

Response	Number (%)
5 (Highly influential)	5 (14%)
4	15 (43%)
3	8 (23%)
2	4 (11%)
1 Not influential at all	3 (9%)

4.7 Which of the following factors affected (either positively or negatively) your decision to do Young Journalism related activities outside the eight Young Journalist Academy teaching days? Please rate each of the factors from 1-5 where 1=not influential at all, and 5=highly influential.

Factor: Own interest in journalism/media

Response	Number (%)
5 Highly influential	4 (11%)
4	10 (29%)
3	9 (26%)
2	3 (9%)
1 Not influential at all	9 (26%)

4.8 Which of the following factors affected (either positively or negatively) your decision to do Young Journalism related activities outside the eight Young Journalist Academy teaching days? Please rate each of the factors from 1-5 where 1=not influential at all, and 5=highly influential.

Factor: Support offered by Young Journalist Academy mentor/team to do so

Response	Number (%)
5 Highly influential	4 (11%)
4	7 (20%)
3	13 (37%)
2	6 (17%)
1 Not influential at all	5 (14%)

4.9 Which of the following factors affected (either positively or negatively) your decision to do Young Journalism related activities outside the eight Young Journalist Academy teaching days? Please rate each of the factors from 1-5 where 1=not influential at all, and 5=highly influential.

Factor: Support provided by my school's SLT to do so

Response	Number (%)

5 Highly influential	5 (14%)
4	7 (20%)
3	11 (31%)
2	5 (14%)
1 Not influential at all	7 (20%)

5.1 To what extent do you agree or disagree with the following statements: The required space was available to deliver the programme effectively at my school

Response	Number (%)
Strongly agree	17 (49%)
Agree	13 (37%)
Neither agree nor disagree	3 (9%)
Disagree	2 (6%)
Strongly disagree	0 (0%)
Don't know	0 (0%)

5.2 To what extent do you agree or disagree with the following statements: The required technology was available to deliver the programme effectively at my school

Response	Number (%)
Strongly agree	11 (31%)
Agree	14 (40%)
Neither agree nor disagree	3 (9%)
Disagree	5 (14%)

| Strongly disagree | 2 (6%) |
| Don't know | 0 (0%) |

6.1 To what extent do you agree or disagree with the following statements: My pupils found the Young Journalist Academy programme to be engaging

Response	Number (%)
Strongly agree	16 (46%)
Agree	8 (23%)
Neither agree nor disagree	9 (26%)
Disagree	2 (6%)
Strongly disagree	0 (0%)
Don't know	0 (0%)

6.2 Which of the following, if any, do you think would help to engage your pupils:

Response	Number (%) selected 'yes'
Changing the way the programme is delivered (i.e. using less PowerPoint)	9 (26%)
Spending time on a different activity (e.g. more article writing/radio production/TV production)	6 (17%)
Having more technology available (e.g. laptops, recording equipment)	5 (14%)
Holding more sessions with the Young Journalist Academy mentor	2 (6%)
Doing more Young Journalist Academy-related activities outside the 8-day programme	2 (6%)
Other (please specify): - Working in smaller groups - Completing project in 6-8 weeks	 1 (3%) 1 (3%)

N.B. Total percentage is less than 100% because some respondents did not select an answer

7.1 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' <u>ideation (ability to come up with ideas)</u>?

Response	Number (%)
Very positive impact	8 (23%)
Somewhat positive impact	14 (40%)
Limited impact	12 (34%)
Somewhat negative impact	0 (0%)
Very negative impact	1 (3%)

7.2 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' <u>behaviour</u>?

Response	Number (%)
Very positive impact	4 (11%)
Somewhat positive impact	6 (17%)
Limited impact	24 (69%)
Somewhat negative impact	1 (3%)
Very negative impact	0 (0%)

7.3 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' <u>reading</u>?

Response	Number (%)
Very positive impact	6 (17%)
Somewhat positive impact	14 (40%)

Limited impact	15 (43%)
Somewhat negative impact	0 (0%)
Very negative impact	0 (0%)

7.4 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' <u>writing skills</u>?

Response	Number (%)
Very positive impact	9 (26%)
Somewhat positive impact	15 (43%)
Limited impact	11 (31%)
Somewhat negative impact	0 (0%)
Very negative impact	0 (0%)

7.5 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' <u>communication skills</u>?

Response	Number (%)
Very positive impact	7 (20%)
Somewhat positive impact	19 (54%)
Limited impact	9 (26%)
Somewhat negative impact	0 (0%)
Very negative impact	0 (0%)

7.6 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' <u>social skills</u>?

Response	Number (%)
Very positive impact	6 (17%)
Somewhat positive impact	20 (57%)
Limited impact	9 (26%)
Somewhat negative impact	0 (0%)
Very negative impact	0 (0%)

7.7 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' creativity?

Response	Number (%)
Very positive impact	7 (20%)
Somewhat positive impact	17 (49%)
Limited impact	11 (31%)
Somewhat negative impact	0 (0%)
Very negative impact	0 (0%)

7.8 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' engagement with culture and the wider world?

Response	Number (%)
Very positive impact	13 (37%)
Somewhat positive impact	13 (37%)
Limited impact	9 (26%)
Somewhat negative impact	0 (0%)

Very negative impact	0 (0%)

7.9 Thinking about your class as a whole, what kind of impact, if any, do you think the Young Journalist Academy experience has had on pupils' <u>confidence</u>?

Response	Number (%)
Very positive impact	9 (26%)
Somewhat positive impact	18 (51%)
Limited impact	8 (23%)
Somewhat negative impact	0 (0%)
Very negative impact	0 (0%)

8.1 Overall, how would you rate Young Journalist Academy?

Response	Number (%)
Very good	13 (37%)
OK	11 (31%)
Good	9 (26%)
Poor	1 (3%)
Very poor	1 (3%)
Don't know	0 (0%)

Appendix M. Raw control condition survey data

Please note: some percentages may add up to slightly above or below 100% due to rounding

1.1 What is your role at the school?	
Response	**Number (%)**
Class teacher	23 (70%)
Head or deputy head teacher	6 (18%)
Other members of the SLT	2 (6%)
Other (please specify): - English lead and class teacher - Year group lead	1 (3%) 1 (3%)
Teaching assistant	0 (0%)

2.1 In the last academic year, has your class taken part in any lessons or other activities related to media or journalism?	
Response	**Number (%)**
Yes	18 (55%)
No	14 (42%)
Don't know	1 (3%)

2.2 How many times over the course of the last academic year has your class taken part in lessons/activities related to media or journalism?	
Response	**Number (%)**
More than ten times	8 (24%)
Two to five times	7 (21%)
Five to ten times	2 (6%)

Once	1 (3%)

N.B. Total percentage is below 100% because this question was only answered by respondents who selected 'yes 'to the previous question.

2.3 Please tick which of the following activities your class has taken part in this year:	
Response	**Number (%) answered 'Yes'**
Watching the news/news clips	17 (52%)
Producing media content, e.g. writing newspaper articles	17 (52%)
Reading newspaper articles	16 (48%)
Teaching the principles of journalism/media, e.g. style, structure, forms	13 (39%)
Listening to radio programmes	5 (15%)
Critical analysis of media outputs	6 (18%)
Other (please specify):	
- Role-play	1 (3%)
- Creative writing	1 (3%)
- Visit to cinema to see how films are produced	1 (3%)
- Writing, recording and editing news bulletins	1 (3%)
- Teaching using Green Screen	1 (3%)

N.B. Total percentage is greater than 100% because respondents could select more than one answer

Appendix N: Memorandum of Understanding for schools

Young Journalist Academy Evaluation

MEMORANDUM OF UNDERSTANDING

This project is exploring how delivery of the Young Journalist Academy (YJA), which provides schools and young people with access to professional levels of training in media production and journalism including article writing, filmmaking, radio production and photography, improve outcomes in writing for children. Its impact will be evaluated by comparing it with the 'teaching as usual' approach using a randomised controlled trial (RCT).

During this project, you will be contacted by both the **Project Team** (Young Journalist Academy), who are responsible for developing and supporting the project, and by the **Evaluation Team** (University College London [UCL] and Behavioural Insights Team [BIT]), who are carrying out an independent evaluation of its effectiveness.

This memorandum of understanding (MoU) explains what your school's participation in the study will entail.

Randomised Controlled Trial (June 2018 July 2019)

The trial will involve your school being randomly assigned either to participate in the Young Journalist Academy (the intervention group) or to continue as normal (the comparison group).

The Young Journalist Academy will provide intervention schools and young people with access to professional levels of training in media production and journalism including article writing, filmmaking, radio production and photography. We plan to work primarily with Year 5 in 100 schools during 2018–19. 'Intervention 'schools (see below) will work with us to establish a Young Journalist Academy in their school.

Schools are asked to pay £1,250 to register their interest and secure their place as part of the trial; Schools in the intervention group will not be asked for any further payment as the rest of the cost of YJA is funded by the EEF as part of the trial; Schools in the comparison group will receive a refund of their £1,250 on allocation to this group; in addition they will receive a payment of X at the end of the trial once all commitments above had been completed.

The following information and evaluation data will be required by the Evaluation and Project teams:

Prior to randomisation

Schools will:

- Provide school URN and LAESTAB number.

- Provide contact details of the Year 5 Lead Project Teacher (valid email address and telephone number) to the Project Team for use by both the Project Team and the Evaluation team.

- Provide, via the means specified by the evaluation team, pupil names, DOB and Unique Pupil Numbers (UPNs) of the Year 5 Project Class, along with details of any setting or streaming by attainment, to the Evaluation Team by the end of March 2018.

Follow the secure procedures requested by Evaluation team to allow parents to object to use of their children's data as part of the evaluation part of this research.

During the evaluation

Participating teachers will:

- Update UPNs and pupil names of Year 5 Project Class by the end of September 2018.

- Facilitate the delivery of Young Journalist Academy in the school, which includes organising training sessions;

- Facilitate visits by the Project Team and/or the Evaluation Team to gather data on the implementation of the intervention or on routine teaching (for example, through observations or interviews etc.).

Facilitate an endof-project written assessment which will be administered by the Evaluation Team.

Use of Data

All pupil data will be treated with the strictest confidence and will be stored in accordance with the Data Protection Act (1998) and with the forthcoming General Data Protection Regulation. Named data will be matched with the National Pupil Database using pupils 'UPNs by the Evaluation Team and shared (anonymously) with the Education

Endowment Foundation. All results will be anonymised so that no schools will be identifiable in the report or dissemination of any results.

For the purpose of research, the pupil data will be linked with information from the National Pupil Database held by the Department for Education, other official records, and shared with the Department for Education, EEF, EEF's data contractor FFT Education and in an anonymised form to the UK Data Archive and for research purposes. Confidentiality will be maintained at all times.

Requirements for Schools

- The school is not participating in another research project or evaluation that would interfere with development and evaluation of the above approach in Year 5.

- The Year 5 Project Teacher will be working with a Year 5 class in the academic year 20182019.

- The school will deliver letters to parents giving them information about the study and an opportunity to object to data processing as part of this evaluation. They will inform the Evaluation Team of any responses arising.

- The school will provide data requested to the Project Team and the Evaluation Team as detailed above. The school will permit the publication of anonymised data collected and its use in presentations.

- Teachers will, at the earliest opportunity, notify the Project Team if there are support or operational issues which could affect their participation.

If the school has to withdraw from the project for operational or other unavoidable reasons, it will notify the UCL/BIT Evaluation Team straight away and, wherever possible, still provide test data for the project.

Responsibilities of the Project Team:

- Provide YJA training

- Act as the first point of contact for any questions about the project

- Provide ongoing support to the school

- Provide information sheets for parents

- Collect Year 5 Project Teacher contact names and email details

Disseminate the research findings.

Responsibilities of the Evaluation Team:

- Conduct the random allocation

- Work with the project team and schools to collect class- and pupil-level data (including name, date of birth, UPN)

- Collect data about how the schools are implementing the intervention, or teaching as usual (control groups)

- Request NPD data using pupil details

- Administer the writing assessment

- Analyse the data from the project

Disseminate the research findings

Appendix O: WAM Prompt

Adapted from Dunsmuir et al. (2015)

Your name: _____

Imagine that you are going to go on a school trip with your class, your teacher, and some other adults. Imagine you could go anywhere at all.

Write a short story about your chosen trip, and some of the things that you did while on the trip.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Appendix P: WSEM Prompt

Adapted from Bruning et al. (2013)

✏️ Thinking about myself as a Writer ✏️

Name: _____

Instructions: Put a tick in **one** box for each question below to show how confident **YOU** feel about the different things a writer does. This isn't a test. We want to know how you really feel when you are writing.

	I'm sure I can't do it	I don't think I can do it	I'm not sure if I can do it or not	I can mostly do it	I'm very sure I can do it
1. I can think of lots of new ideas for my writing.					
2. I can use my ideas in my writing.					
3. I can think of the words I need to write down my ideas.					
4. I can show how I feel in my writing					
5. I know where to place my ideas in my writing.					
6. I can spell my words correctly.					

7. I can write whole sentences.					

Please turn over to the other side

	I'm sure I can't do it	I don't think I can do it	I'm not sure if I can do it or not	I can mostly do it	I'm very sure I can do it
8. I can use punctuation marks (like full stops and question marks) correctly.					
9. I can begin my paragraphs in the right places.					
10. I can focus on my writing for at least half an hour.					
11. I can start writing quickly.					
12. I am always calm and in control when I write.					
13. I can think of my goals for my writing before I write.					

14. I can keep on writing even when it gets difficult.					
15. I enjoy writing.					
16. I am confident as a writer					

Thank you for helping us by thinking about writing.

Appendix Q: School and parent information sheets, objection forms, and revised GDPR privacy notice

School information sheet

The EEF/RSA 'Learning about Culture' Trial of

The Young Journalist Academy

What is this about?

Young Journalist Academy are currently taking part in a project funded by EEF and the RSA. This aims to improve our understanding of interventions to improve pupils writing skills, including Young Journalist Academy. The effectiveness of the project will be evaluated by a team from University College London and the Behavioural Insights Team. **This is a great opportunity to help us understand the impact of Young Journalist Academy and potentially receive the programme at a significantly reduced cost!**

What will the project look like?

The project is investigating the effect of the Young Journalist Academy programme (the 'project team') which will work in schools to deliver the programme. The Young Journalist Academy provides schools and young people with access to professional levels of training in media production and journalism including article writing, filmmaking, radio production and photography.

We plan to work primarily with Year 5 in 100 schools during 2018–19. 'Intervention' schools (see below) will work with us to establish a Young Journalist Academy in their school. You can find out more about the Young Journalist Academy at www.youngjournalistacademy.com

What are 'intervention' and 'comparison' schools?

An important element of these EEF projects is that schools interested in participating are randomly allocated to either the intervention group, who will receive the training, or a comparison group, who contribute to the data required for comparison. It is important that all interested schools understand that they could be in either group. ***It is important that you understand you could be a comparison school.***

What commitment would this project require?

If you are an 'intervention' school, you would need to commit to allowing the Year 5 project teacher to:

➢ provide student profile and attainment data and UPNs;

➢ facilitate the delivery of Young Journalist Academy in the school, which includes organising training sessions;

➢ allow the research team access to collect data (for example through observations and interviews);

➢ allow the evaluation team to visit your school to administer a writing assessment at the end of the project.

If you are a 'comparison' school, you would need to commit to allowing the project teacher to:

➢ provide student profile and attainment data and UPNs;

➢ allow the evaluation team to visit your school to administer a writing assessment at the end of the project.

Finances

- Schools are asked to pay £1,250 to register their interest and secure their place as part of the trial

- Schools in the **intervention** group will not be asked for any further payment as the rest of the cost of YJA is funded by the EEF as part of the trial;

Schools in the **comparison** group will receive a refund of their £1,250 on allocation to this group; in addition they will receive a payment of X at the end of the trial once all commitments above had been completed.

If you would like to know more, or if you have any questions, please contact Rob Pitman at Paradigm Arts by email at rob@paradigmarts.co.uk or by phone on 07885 384 667.

Parent information sheet

YOUNG JOURNALIST ACADEMY

What is this about?

The Young Journalist Academy (YJA) is currently working on a project funded by Education Endowment Foundation and the Royal Society of the Arts, exploring whether and how their programme helps to improve outcomes in writing for the children in the schools where they work. The effectiveness of the project will be researched by a team from UCL Institute of Education and the Behavioural Insights Team (the 'evaluation team'). This research has been reviewed and approved by the research ethics committee of UCL Institute of Education.

What will the project look like?

The project is investigating the effect of the Young Journalist Academy programme, run by a team from Paradigm Arts and c1media (the 'project team'), which will work in schools to delivery their usual programme designed to encourage young people to engage with news and current affairs by creating their own School Newsroom, media literacy projects and working across school departments. This opportunity enables schools to work with industry trained broadcast journalists and other professionals to train students and staff in all aspects of broadcast journalism. These include writing, producing, filming and recording.

We plan to work with 100 primary schools, particularly their Year 5 classes during 2018–19. 'Intervention' schools (see below) will participate in Young Journalist Academy's usual programme for a reduced price during this first year. You can find out more about Young Journalist Academy from http://youngjournalistacademy.com/about/.

What are 'intervention' and 'comparison' schools?

An important element of these EEF projects is that schools are randomly chosen either to be in the intervention group, who will receive the training this year, or a comparison group, who contribute to the data required for comparison (and might choose to do the training in future).

Whether your child's teacher will receive the training this year ('intervention' schools) or not ('comparison' schools) will be randomly decided by evaluators from UCL Institute of Education and the Behavioural Insights Team to help them understand how effective YJA has been. If your child's teacher does not receive the training this year, they will receive a payment to allow the school to take part in YJA in future years (or for other purposes, if they prefer).

What does this mean for me as a parent?

As part of measuring the success of this training programme, your child will be asked to complete a writing test and survey during their normal classes towards the end of the school year. This will take about half an hour. We are doing this test for the purposes of the research project, to help us understand if the Young Journalist Academy programme helps children like yours with their writing. We are also asking to obtain your child's Unique Pupil Number (issued by the UK Government's Department of Education) to allow longer term understanding of whether this writing programme worked (for example by accessing their age 11 English SATS test scores). We will then save this information in a data format that will prevent anyone from identifying your child.

This data will be linked with the UK Government's Department for Education's National Pupil Database and shared with the project team, the Department for Education, the Education Endowment Foundation (EEF, who are funding this research), EEF's data contractor FFT Education and kept in an anonymised form in the UK Data Archive. No information that can identify individual children will be made available to anyone outside these teams and your child's school. This data will be kept securely under password protection. We will not use your child's name or the name of the school in any report arising from the research, and no information that could otherwise identify your child will be made public.

Although we think this project will benefit your child and that they will enjoy being part of it, you have the right to ask us not to use your child's data (either writing assessment performance or data from the National Pupil Database) in this way. Please see the letter that came with this information sheet or contact your child's class teacher. If you have any concerns and would like to know more, or if you have any questions, please contact Louise Jones at the Behavioural Insights Team by email at louise.jones@bi.team or by telephone on 07804494899.

Objection letter

Dear Parent / Carer,

Your child's school has applied to take part in research that aims to improve their self-confidence and performance in writing. The Young Journalist Academy programme, run by a team Paradigm Arts and c1 media (the 'project team'), work in schools to encourage young people to engage with news and current affairs by creating their own School Newsroom, media literacy projects and working across school departments.

This process will then be researched by a team from UCL Institute of Education, the Behavioural Insights Team, along with some evaluation by the Royal Society of Arts (the 'evaluation team'). There's more information on the information sheet that came with this letter.

Not all schools in the study will necessarily take part in Young Journalist Academy this year. Whether your child's school will take part this year will be decided by the evaluation team at random to help them understand how effective the project has been.

As part of measuring the success of this training programme, all Year 5 children will complete a classroom based writing test and survey towards the end of the year. This will take about half an hour. We are doing this test is for the purposes of the research project, to help us understand if the Young Journalist Academy programme helps children like yours with their writing.

Your child's name and other data held by the school, alongside their writing test scores, will be collected by the evaluation and project teams. No information that can identify individual children will be made available to anyone outside these teams and your child's school. This data will be kept securely under password protection. We will not use your child's name or the name of the school in any report arising from the research, and no information that could otherwise identify your child will be made public.

We are asking to obtain your child's Unique Pupil Number (issued by the UK Government's Department of Education) to allow us to link up our data with their National Pupil Database and other official records to understand whether being part of this project is linked with improved test scores (such as their age 11 English SATS test scores) when they are older. This involves us sharing data with the Department for Education, the Education Endowment Foundation (EEF, who funded the trial), EEF's data contractor FFT Education and in a form that will prevent anyone from identifying your child to the UK Data Archive.

This research has been reviewed and approved by the research ethics committee of UCL Institute of Education.

If you have questions about any aspect of the planned research and use of data that you would like to ask before replying, please contact the research team via Louise Jones at the Behavioural Insights Team by email at louise.jones@bi.team or by phone on 07804494899.

Because we are doing this research to improve understanding about what works in improving pupils' education, **if you are happy for information about your child to be used in the Young Journalist Academy research project you do not need to do anything.** Thank you for your help with this research, your support is much appreciated.

If you **DO NOT** want your information about your child (either their writing test scores or or data from the National Pupil Database) to be used to understand whether the Young Journalist Academy programme can help children to write better, please complete the enclosed form and return it to your child's school by [INSERT DATE]. If you do this then no information about your child will be shared with the evaluation or project teams at any point during the project. This does not affect your child's ability to take part in any Young Journalist Academy activities that happen in their school, if they wish to do so.

Young Journalist Academy research programme

(If you are happy for your child to participate in the research on whether this programme improves writing, you DO NOT need to return this form.)

I DO NOT wish my data about my child to be collected as part of this research.

Child's name: …………………………………………………Date of birth: ………………

Child's class Teacher: …………………………………………………………………..

School:…………………………………………………………………………………

Parent name (BLOCK CAPITALS) …………………………………………………

Parent signature: ……………………………………………………………………

Date ……………………………………………

(Please return the completed form to your child's class teacher.)

Revised GDPR privacy information

EEF/RSA EVALUATION OF YOUNG JOURNALIST ACADEMY

Data Privacy Notice

Dear Parent,

We've previously been in touch because your school is taking part in a project funded by the Education Endowment Foundation (EEF) to understand the potential benefits of Young Journalist Academy. As part of that information, you were given the opportunity to tell your school not to pass any data about your child to us to be used as part of this project. **Please rest assured that if you contacted the school to make this request, nothing in this letter changes that.**

We wanted to get in touch again to provide you with further details about the way will be handling pupils 'data as part of this project. **It is very important to us that that we do this responsibly and providing these details are an important part of that.** They are also important in fulfilling our responsibilities under the UK's data protection laws, which we take very seriously. These require us to provide you with some specific information about **our plans** and **your rights**.

Some of this description involves rather technical terms, which we've left in so you know the official concepts we are talking about. We've tried to keep the explanations as simple as possible. If we haven't managed that well enough and you have any questions now, or at any point during this project, then you should contact a member of the team with the first point of call being XXX.

Our Plans

- Using pupils' data as part of research is not something we do without thinking about it. Under data protection law, we require a 'lawful basis' for the data processing that we carry out. UCL will be using the lawful basis known as the 'public task' basis, while the lawful basis BIT are using is known as the 'legitimate interests' basis (it is different at UCL and BIT because UCL is a university). To use the 'legitimate interests' basis, we must consider why this is a legitimate interest and inform you of this. Here, it is because our work is for the purpose of promoting the education or well-being of children in England and couldn't be achieved without analysing these test scores. We balance that against your and child's rights by providing you with the right to object to our use of your child's data in this way.

- Your child's name and other data held by the school, alongside the test scores we will collect, will be collected and processed by us for the purpose of understanding how participating in this project has affected their learning. No information that can identify individual children will be made available to anyone outside these teams and your child's school (with an important exception, as explained below). We will also obtain your child's UPN (Unique Pupil Number) from the school to allow us to link up our data with the National Pupil Database (held by the Department for Education) and other official records to understand whether being part of this project is linked with test scores when they are older. This involves us sharing data with the Department for Education (part of the UK Government), the Education Endowment Foundation (who funded the trial), EEF's data processor FFT Education and (in a form that will prevent anyone from identifying your child) to the UK Data Archive.

- The data we hold will be kept securely at all times, transferred using secure (encrypted) methods, and kept on secure computer systems at UCL and BIT's offices under password protection. We will never use your child's name or the name of the school in any report arising from the research, and no information that could otherwise identify your child will be made public.

- Pupils' personal data will be processed by us only for the purposes of this research project. Once that is complete then the data will be securely destroyed from our computer systems. Personal data will certainly not be more than 10 years, in line with UCL's policy on storing research data.

Your rights

- Under data protection law, you have a right to be informed about our plans. This letter, as well as the information that you previously received from your school, are all part of this. You also need to know exactly who is involved in the data processing. In legal terms, University College London (UCL) and the Behavioural Insights Team (BIT) are considered joint data controllers for this project. The law requires our organisations to have named Data Protection Officers, who are ultimately responsible for overseeing data processing that goes on in their respective organisations. UCL's Data Protection Officer is Lee Shailer, who can be contacted via data-protection@ucl.ac.uk. BIT's Data Protection Officer is Lizetta Lyster, who can be contacted via lizetta.lyster@bi.team. You should contact these individuals if you have any complaints about how we are processing data. However, if you remain unsatisfied, you may wish to contact the Information Commissioner's

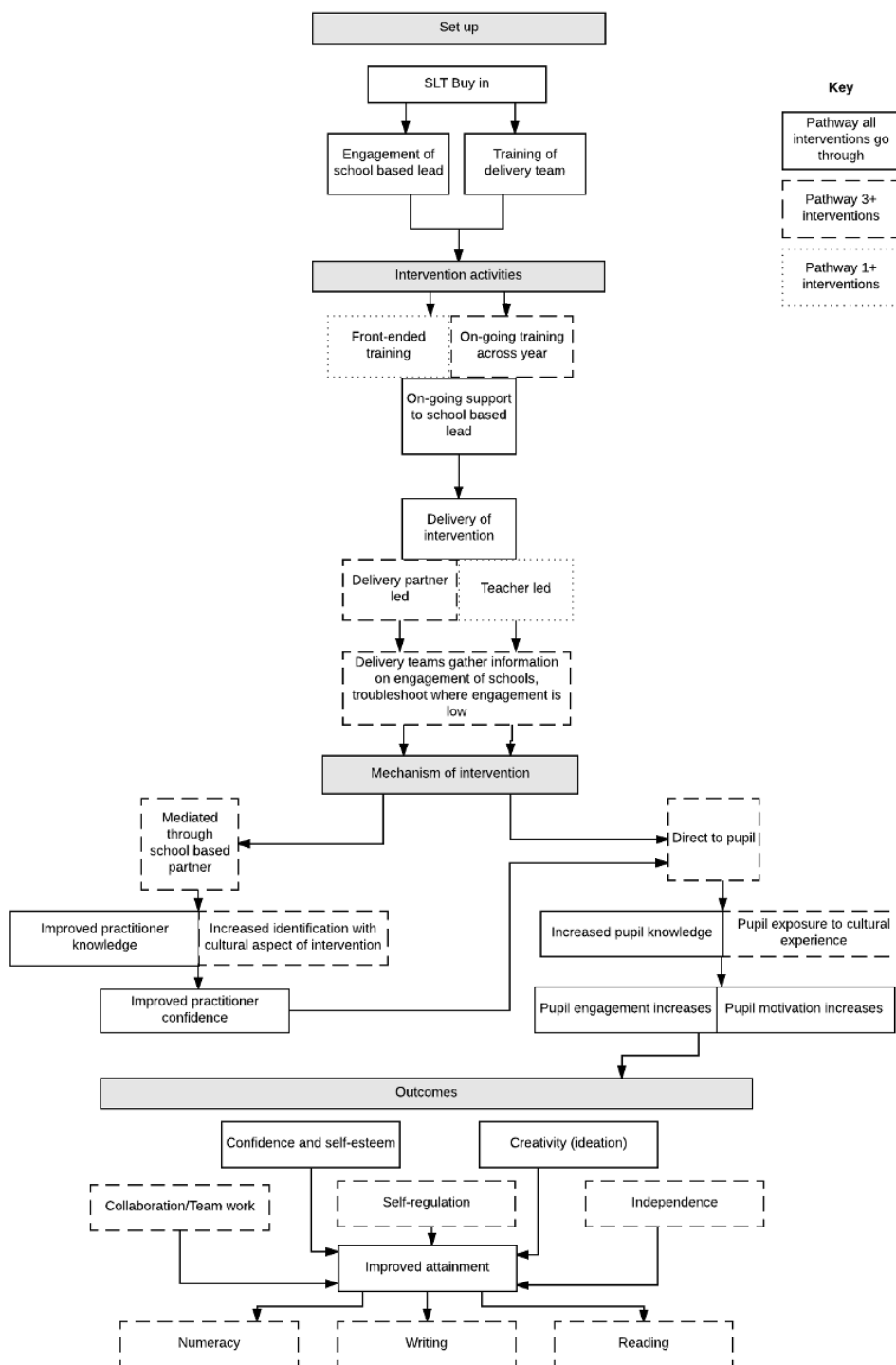Office (ICO). Contact details, and details of data subject rights, are available on the ICO website at:
https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/individuals-rights/

As noted above, we provided you with the right to object to data processing before schools handed any information over to us. You can also contact us at any point during the project to request information we hold about your child, to request rectification of any information that is incorrect, to stop using their data as part of the project or to destroy their data. If you wish to make such a request or ask any questions about it then please contact us. The best place to start is to contact XXX.

Once again, we are extremely grateful to you for supporting this project. We hope to learn a lot about the role cultural and arts education can play in supporting pupils 'learning.

Yours sincerely,

Appendix R: Amalgamated logic model for all five Learning about Culture trials