Grammar for Writing
University of York
Louise Tracey

Education Endowment Foundation

Evaluation Summary

Age range	Year 6
Number of pupils	Approx. 7,500
Number of schools	150 schools
Design	School-level effectiveness randomized controlled trial
Primary Outcome	Key Stage 2 Writing

The Intervention

Grammar for Writing is a way of teaching writing that assumes that rather than teaching grammatical rules in the abstract, teachers should help pupils to understand how linguistic structures convey meaning. Consequently, the programme aims to improve writing by developing pupils' understanding of grammatical choices. Underpinned by key pedagogical principles Grammar for Writing is embedded in the context of teaching about writing genres. The core elements of the programme encompass the use of grammar terms, linking grammar effects in writing, and using talk to develop discussion about choices and effects The programme is designed to be delivered by teachers as standalone units of work or as a series of units within a whole class setting. Each unit is around 4 weeks' worth of work.

The concept of improving children's grammar in parallel with their writing by using a contextual approach is a promising idea that, as yet, lacks conclusive evidence in the primary school phase. There have been several, developer-led trials in secondary schools which have demonstrated positive results (Jones et al, 2012; Myhill et al, 2012). An efficacy trial funded by the Education Endowment Foundation (Torgerson et al, 2014) looked at whole-class and small group delivery in a 4-week version of the programme adapted for Year 6 after Key Stage 2 SaTs assessments. However, this found only limited effects measured by children's performance on the GL Progress in English assessment. For the whole-class intervention, there was a small and statistically non-significant effect. The impact for those additionally taught Grammar for Writing in small groups was higher, although it was not much higher than those taught in small groups without Grammar for Writing. This evaluation of Grammar for Writing focuses on whole-class delivery of two units of Grammar for Writing during Year 6.

Methods

Research questions

The primary research question is:

- How effective is Grammar for Writing in improving the writing skills in Year 6 pupils.

The secondary research question addresses the impact of Grammar for Writing on overall literacy attainment. An intermediary hypothesis relates to the impact of teacher grammar knowledge on pupil outcomes and the impact of Grammar for Writing on this teacher knowledge.

Research Design

This will be two-arm effectiveness RCT with randomization occurring at the school-level to reduce the possibilities of diffusion which could occur with an in-school design. For the purposes of this trial two units of work will be delivered within a whole-class context: narrative writing and persuasive writing. There will be 4 days of CPD provided to teachers through the school year, delivered by the University of Exeter and Babcock LDP. This CPD will include the provision of teaching materials for the two units of work.

Participants

Eligible schools will not have taken part in the previous Grammar for Writing trial or have implemented the programme previously. Although they do not have to be two-form entry, very small schools (fewer than 20 Year 6 pupils) will be kept to a minimum.

Half of schools will be from the North East and the other half will come from across the rest of England. There will be a high proportion of disadvantaged schools (on average 29% EverFSM as defined in the National Pupil Database).

Recruitment will be conducted by the project team (University of Exeter), with support from the evaluation team (University of York).

Randomisation

Schools will only be eligible for randomisation after:

- Signing an Memorandum of Understanding; and
- Providing pre-test data requested in the Memorandum of Understanding (including pupil UPNs, providing teacher contact details and completion of teacher pre-intervention survey (see process evaluation section below).

Randomisation will be conducted at the school level using minimisation. Minimisation uses algorithms to ensure balance at baseline and permits ongoing allocation so schools know which condition they have been assigned to soon after recruitment. Schools will be stratified by region (North East / not-North East). Randomisation will be conducted and recorded by the evaluation team using MinimPy software (MinimPy, 2013; Saghaei & Saghaei, 2011).

Incentives

Intervention schools will receive the Grammar for Writing programme and associated materials and training at the reduced rate of £500. Control schools will receive £500 on completion of all requested measures, at the end of the intervention period. This can then be used towards funding Grammar for Writing training if desired. This reflects the fact that the burden placed on schools is not high and avoids potential problems such as ethical issues if the intervention was not shown to be effective. In addition teachers will receive an extra payment of £20 in vouchers in exchange for completing the pre- and post-intervention on-line surveys. For both intervention and control schools, regular communication (eg an evaluation newsletter) will remind them of evaluation plans and maintain engagement.

Outcome Measures

The primary outcome will be the combined results of two tasks from past Key Stage 2 (KS2) writing assessments (pre-2013). Past Key Stage 2 assessments have been chosen because the current writing assessment for KS2 consists of a portfolio of teacher-assessed work which, whilst externally moderated, is judged to be 'working toward', 'working at' or 'working above' the expected standard for the end of Key Stage 2. The advantage of using past KS2 writing tasks is that they can be administered in controlled conditions within schools and have a set marking scheme which is sufficiently graded to be able to conduct a meaningful and sufficiently robust analysis to assess the impact of the programme on KS2 writing. The tasks will be selected by the evaluation team to include one longer written task and one shorter written task, covering both persuasive and narrative writing. The project team will remain blind to their exact content. The assessments will be administered in schools independently by the National Foundation for Educational Research (NfER) to ensure controlled conditions and to reduce any burden on schools. They will be marked by a team of assessors, unaware of allocation, at the University of York. As narrative and persuasive writing are implicit within the KS2 curriculum, this primary measure will not be inherent to treatment – teachers in the control condition will also be teaching their pupils to write narratively and persuasively.

It is also important to assess whether any improvement in these aspects of writing have been at the expense of other elements of literacy, maybe as a result of reduced focus on these. For this reason, secondary outcomes include KS2 scores on each element of literacy (writing; reading; grammar, punctuation and spelling) separately and together for the same pupils. Using primarily nationally collected data will minimize costs and the burden on schools and pupils. These measures are high in contextual validity and, since they constitute the main indicators of school and pupil academic performance, all teachers (intervention and control) will be focused on ensuring that pupils succeed on them. With the addition of the past KS2 writing tasks, the proposed outcome measures will provide a measure of all-round performance on literacy, and, specifically, the legacy effect on writing.

The primary pre-test measure will be the Key Stage 1 writing results (obtained from the National Pupil Database). The KS1 English results were highly correlated with the previous KS2 assessments in English (0.73) and we assume that this remains high using the KS2 and KS1 writing measures proposed (EEF, 2013).

An intermediary measure will be the teacher 'grammar quiz' developed for inclusion in the pre- and post-test teacher survey (see Process evaluation below). The pre-test 'quiz' will be that developed by the Exeter team for use as part of the 'Grammar for Writing' training. As the pre-test will be taken prior to allocation teachers' and researchers' will be blind to allocation. A similar quiz will be developed for the post-intervention survey either by the Exeter Team or, if this is not

deemed appropriate (ie it is inherent to treatment), by the evaluation team. The final decision for this will be made by February 2017 and the protocol updated to reflect the decision.

Sample size calculations

The statistical power of the proposed analyses was estimated using Optimal Design software, with assumptions as follows:

Students per school per class: 25 (i.e. 50 per treatment per school)
Pre-post correlation (squared): +0.53
Intraclass correlation: 0.21
Criterion for statistical significance: $p < .05$
Power: 0.80

A sample of 150 schools would be sufficient to detect an effect size of 0.16. This is based on the assumption of a similar pre-/post correlation as the previous KS1/KS2 assessments.

Assuming 16 FSM pupils per school, this sample of 150 schools would enable an effect size of 0.18 to be detected in the FSM sub-sample.

Analysis plan

The impact evaluation will use Hierarchical Linear Analysis (HLM), a multilevel analysis in which pupils are nested within schools. This makes it possible to separate within school variation in the outcome from between school variation. The analyses will use an intent-to-treat-design which means that outcomes data will be treated in the main analysis according to the condition allocated (control or intervention), not that actually received. The pre-test will be used as the covariate. Effect sizes will be calculated using Hedge's g. In addition to analysing pupil outcomes teacher outcomes will be analysed as a secondary outcome to assess the impact of the programme on teacher grammar knowledge.

After the main analyses including all pupils, subgroup analyses will be carried out for pupils eligible for FSM; boys and girls; high and low achievers on the pre-test; and high and low implementation fidelity. A mediator analysis will also be undertaken to account for the impact of the possible change in teacher knowledge as a result of the intervention on pupil knowledge.

Implementation and process evaluation methods

Previous evaluations have indicated the importance of fidelity of implementation on outcomes (Eames et al, 2009). The previous EEF evaluation of Grammar for Writing was a 'light touch' approach, with fidelity scores established using a measure developed by the implementation team. It did, however, indicate high levels of fidelity to the programme but low (and non-significant) impact on outcomes (Torgerson et al. 2014). To effectively evaluate this intense intervention a thorough process evaluation is necessary.

Our proposed process evaluation aims to:
- examine more closely the relationship between the level of implementation of the intervention and its impact on pupil outcomes;
- explain variability in implementation, including understanding the context of the implementation and social processes within schools; and
- address possible barriers to implementation.

We aim to do so through structured observations, teacher surveys, and interviews with selected teachers and school literacy coordinators.

a) Structured observations – A fidelity measure will be used encompassing the three core components of the programme (use of grammar terms, linking grammar effects in writing, and using talk to develop discussion about choices and effects). The proposed checklist will also reflect the more detailed dynamics of the programme, other writing/grammar techniques taught in Year 6 and the wider classroom context eg, levels of pupil engagement. It is intended that this implementation fidelity check will provide a quantitative measure to inform the analysis of pupil outcomes. Teachers would be categorised as high, medium or low implementers and the effects of their level of implementation correlated with their pupils' outcomes.

A random subsample of 15 schools will be observed, include a small number of control schools. This would enable us to compare the conditions and identify any common strands in lesson delivery and content. The observation schedule will be developed in conjunction with the project team and be used by both the project and evaluation teams (intervention schools only for the project team). Observers will also attend a Grammar for Writing training session to ground their observations of the intervention.

b) Teacher surveys – All teachers would be asked to complete a Likert-style on-line survey before and after the intervention using Qualtrics (Qualtrics, 2015). This will provide information regarding their professional and academic background, linguistic subject knowledge, confidence in teaching literacy, specifically grammar and writing, any schemes of work used in their literacy teaching and contextual classroom factors. This will include a short 'grammar quiz'. Post-test intervention teachers will also be asked about adherence to the programme, the training and materials received, attitudes towards, and confidence in, Grammar for Writing. Emails and reminder phone calls will be used to encourage completion.

c) Semi-structured interviews with selected teachers – The teacher interviews would primarily unpack the questions asked in the survey in more detail. In addition the teachers would be asked to discuss the classroom observation and the classroom context, in particular any adaptations made by the teacher to accommodate their pupils' needs. Ideally these interviews would take place after the classroom observations to enable greater understanding of the lesson observed. Teachers in the intervention condition would be asked additional questions, including, for example, the acceptability of the programme, and any barriers to implementation.

d) Semi-structured interviews with selected literacy co-ordinators – The literacy co-ordinators in the observed schools will be briefly interviewed to understand school demographics, needs and context, literacy priorities and any challenges in meeting those needs.

All interview data would be analysed using NVivo 10 (QSR International, 2012) which would enable researchers to quantify the contextual data provided.

To support this process evaluation we will analyse routine data collected by the Exeter team. This will include fidelity data collected during CPD days. The project and evaluation teams will work closely together to minimise burden on schools and a Data Sharing Agreement will be established for the exchange of information and research data collected (where appropriate). The observation measures will be piloted both in schools currently using the Grammar for Writing programme and

those not implementing the programme. The survey and interview schedules would be shared with a small number of teaching professionals, including staff in these schools, to judge the acceptability and accessibility of the measures.

Process evaluation data will be triangulated to ensure the robustness of any research findings, including an in-depth understanding of the programme as intended and as implemented within a natural setting and to identify (1) how schools can support staff in implementing Grammar for Writing, and (2) any appropriate adaptations of the programme to assist with any further development and possible scale-up of the intervention.

Costs

Following EEF guidelines, the evaluation team will provide a cost per pupil per year for the first year and an estimation of the first three years of the intervention (following years will be less expensive due to non-recurrent costs being excluded, eg, lower training costs in subsequent years to cover new staff or top-up costs only). Costs of implementation will be systematically identified and are likely to include training days (paying trainers, cost of materials, printing, venue, refreshments, travel) and resources (photocopying costs for Grammar for Writing activities). Opportunity, rather than financial costs, will be split out separately, eg, teacher time for training and planning. Cover costs for training will also be identified separately because some schools will pay for a supply teacher whereas others will manage by reallocating existing staff. Cost implications will be identified through a pro forma spreadsheet, discussions with the project team and teacher feedback from the survey and school visits. If possible, comparison of Grammar for Writing schools with control schools would identify any extra costs or cost savings associated with the intervention.

Ethics and registration

Informed ('opt-in') consent will be obtained from Head teachers and teachers. Opt-out consent will be offered to parents to withdraw their child's data in the evaluation. Consent includes linking to the National Pupil Database and data archiving. Data will be handled in accordance with the Data Protection Act (1998).

Ethical Approval has been obtained from:
- Education Ethics Committee, University of York (Ref: 16/18)
- Social Sciences and International Studies Ethics Committee, University of Exeter

This trial has been registered at http://www.isrctn.com/ (to be completed).

Personnel

Evaluation Team

Dr Louise Tracey (Principal Investigator), Research Fellow, Psychology in Education Research Centre, Department of Education, University of York.

Dr Jan R. Boehnke (Co-Investigator), Hull York Medical School and Department of Health Sciences, University of York

Louise Elliott (Co-Investigator), Research Fellow, York Trials Unit, Department of Health Sciences, University of York.

Dr Claudine Bowyer-Crane (Co-Investigator), Senior Lecturer, Psychology in Education Research Centre, Department of Education, University of York.

The evaluation team will be responsible for the design, randomization, data collection, analysis and reporting of the evaluation.

Project Team

Professor Debra Myhill, Director of the Centre for Research in Writing, Graduate School of Education, University of Exeter.

Dr Susan Jones, Senior Lecturer, Graduate School of Education, University of Exeter.

Dr Helen Lines, Research Fellow, Graduate School of Education, University of Exeter.

Sara Venner, Research Fellow, Graduate School of Education, University of Exeter.

The project team will be responsible for school recruitment, intervention development, training and delivery of the programme.

Risks

Risk	Preventative measures	Likelihood
Insufficient schools recruited	• emphasise intervention not very disruptive on the curriculum • emphasise evaluation not very onerous • primary post-test is KS2 SATs to minimise burden • promising results from the Myhill et al, 2012 secondary school trial as persuasive tool	medium to high
High attrition from evaluation, especially of controls	• initial recruitment conference to explain RCT and value of controls • regular "newsletter"/contact • financial incentive to controls • financial incentive to individual teachers for completing measures • over-recruit required number of schools	low
High attrition from Grammar for Writing or poor implementation	• responsive trainers with good monitoring systems • previous trial had 4/55 ie only 7% attrition at school level	low
Poor/inconsistent training of teachers by trainers	• the project team have prior experience through secondary school programme and previous trial • rollout provided by University of Exeter team, and scaling made possible through collaboration with Babcock LDP	low
Staff turnover	• system for notification of teacher turnover or sickness so new staff get trained quickly	medium

Timeline

	School Year 2015-2016			School Year 2016-2017						School Year 2017-2018		
	Spring Term	Summer Term		Autumn Term		Spring Term		Summer Term		Autumn Term		Spring Term
Recruit schools	X	X	X									
Pilot measures with non-study schools			X									
Randomise schools			X									
Collect pre-test data from schools		X	X									
Confirm teachers / class lists				X								
Training and delivery				X	X	X	X	X	X			
Classroom observations and interviews						X	X					
Collect post-test data								X	X			
Conduct Analyses								X	X	X	X	
Main report											X	X

8

References

Education Endowment Foundation (EEF) (2013). *Pre-testing in EEF evaluations*. Accessed at: https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Pre-testing_paper.pdf, 17 January, 2017.

Jones, S.M., Myhill, D.A., Bailey, T.C. (2013). Grammar for Writing? an investigation into the effect of Contextualised Grammar Teaching on Student Writing. *Reading and Writing*, 26(8), 1241-1263.

Myhill, D.A., Jones, S.M., Lines, H., Watson, A. (2012). Re-thinking grammar: the impact of embedded grammar teaching on students' writing and students' metalinguistic understanding. *Research Papers in Education*, 27(2), 139-166.

Saghaei, M. and Saghaei, S. (2011) Implementation of an open-source customizable minimization program for allocation of patients to parallel groups in clinical trials. *Journal of Biomedical Science and Engineering*, 4, 734-739

Torgerson, D., Torgerson, C., Mitchell, N., Buckley, H., Ainsworth, H., Heaps, C. & Jefferson, L. (2014). *Grammar for Writing. Evaluation Report and Executive Summary.* London: Education Endowment Foundation.