

**GRAMMAR FOR WRITING?**  
**AN INVESTIGATION OF THE EFFECTS OF CONTEXTUALISED GRAMMAR TEACHING ON**  
**STUDENTS' WRITING.**

Susan Jones *University of Exeter UK*

Debra Myhill *University of Exeter UK; Professorial Fellow University of Wollongong, Australia*

Trevor Bailey *University of Exeter UK*

**Address for correspondence:**

Debra Myhill

University of Exeter

Graduate School of Education

Heavitree Road

Exeter EX1 2LU

Email: [d.a.myhill@ex.ac.uk](mailto:d.a.myhill@ex.ac.uk)

Tel: +44 1392 724767

**Abstract**

The role of grammar instruction in the teaching of writing is contested in most Anglophone countries, with several robust meta-analyses finding no evidence of any beneficial effect. However, existing research is limited in that it only considers isolated grammar instruction and offers no theorisation of an instructional relationship between grammar and writing. This study, drawing on a theorised understanding of grammar as a meaning-making resource for writing development, set out to investigate the impact of contextualised grammar instruction on students' writing performance. The study adopted a mixed-methods approach, with a randomised controlled trial and a complementary qualitative study. The statistical analyses indicate a positive effect on writing performance for the intervention group ( $e = 0.21$ ;  $p < 0.001$ ); but the study also indicates that the intervention impact differentially on different sub-groups, benefiting able writers more than weaker writers. The study is significant in being the first to supply rigorous, theorised evidence for the potential benefits of teaching grammar to support development in writing.

**Keywords:** grammar; linguistics; writing; subject knowledge.

## **1.0 Introduction**

The instructional benefit of teaching of grammar in first language English curricula is contested in both research and professional literature in Anglophone countries (Gordon 2005; Keith 1997; Wyse 2004) and in the Netherlands (van Gelderen & Oostdam 2005; van Gelderen 2006). Indeed, in the 1960s and 1970s, following the Dartmouth Conference (Dixon 1975), most Anglophone countries (for example, England, Australia, New Zealand and the USA) abandoned grammar instruction on the grounds that it was ineffectual in supporting language development, particularly writing development (Locke 2009). More recently, driven principally by policy imperatives, grammar has been re-introduced into the English curriculum in England, and a parallel process is currently occurring in Australia. However, there is no clarity or agreement about the role of grammar in the English curriculum and it remains a strongly contested issue (Myhill & Jones 2011; Myhill 2011). The uncertain role of grammar in the language curriculum is set within an international context, in Anglophone countries particularly, expressing concerns about the writing attainment of children (NCW 2003; OFSTED 2009; Salah-Din 2008). In England, for example, in 2011, 32% of boys and 19% of girls entering secondary education had not achieved the baseline standard in writing expected for their age group, compared with 20% of boys and 13% of girls who had not achieved the baseline in reading (DfE 2011). However, there have been no systematic studies of whether making meaningful connections between particular linguistic structures and particular writing tasks supports the development of students' writing. This paper reports on the outcomes of a randomised controlled trial, funded by the Economic and Social Research Council, which investigated the impact of contextualised grammar teaching on students' writing performance.

### **1.1 The effectiveness of grammar teaching**

Empirical studies investigating the efficacy of grammar teaching provide little evidence of any beneficial impact upon students' competence in writing. Robust meta-analyses by Braddock, Lloyd-Jones & Schoer (1963), Hillocks (1986) and most recently, by the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI) (Andrews, Torgerson, Beverton, Freeman, Locke, Low, Robinson & Zhu 2006; EPPI 2004) have concluded that there is no evidence that teaching grammar is of benefit in supporting writing development. Indeed, Hillocks and Smith (1991, 602) argue that "research over a period of nearly 90 years has consistently shown that the teaching of school grammar has little or no effect on students".

There are, however, several major difficulties with almost all of the research that these reviews represent. The first is that studies repeatedly investigate whether various forms of grammar teaching, such as learning transformational grammar, grammar exercises and drills, or parsing sentences, improve writing. The emphasis is on teaching grammar

in the hope that it might have an impact on writing outcomes. In many of the studies (for example, Bateman & Zidonis 1966; Elley, Barham, Lamb & Wylie 1975, 1979; Robinson 1959) isolated grammar lessons are taught as part of a curriculum programme in grammar, and the writing measures used to draw empirical conclusions are produced in a different teaching context. Robinson (1959) tested grammatical knowledge and compared this with the quality of their composition - she correlated a grammar test with impression marking, and looked only at word classes. Bateman and Zidonis (1996) taught a transformational grammar course, with the purpose of “determining the effects of a study of transformational-generative grammar on the language growth of secondary school pupils” (Elley et al. 1979, 98). The Elley et al. study (1975, 1979) had three treatment groups: The first was a course typical of English classes at that time in New Zealand using a textbook addressing grammar, comprehension, and writing; the second was a reading-writing course where students spent 40% of their time free reading, 40% sharing a class reader, and 20% writing; and the final treatment group was a transformational grammar course with the intention of helping students “see how they can discover facts about their language and how they use it” (1975,28). Students in this group were taught about such things as sentence combining, subordination, participial modifiers, and deep and surface structures. A second difficulty with the few existing studies is that many are small -scale. The Bateman and Zidonis study, for example, had a sample of 41 students.

A further difficulty is that none of the studies theorise an instructional relationship between grammar and writing, which might inform the design of an appropriate pedagogical approach. The studies are all located in very different educational jurisdictions, with differing pre-existing curricular emphases on grammar. In New Zealand, for example, (Elley et al. 1975; 1979) there was growing unease about the efficacy of traditional grammar teaching amongst educational professionals, but also a back-to-basics call at policy level which appeared to advocate “strong doses of English grammar as a cure for some of our educational ills” (Elley et al. 1975,3). But none directly address the inter-relationship of grammar and writing, or offer a theoretical account of such an inter-relationship.

## **1.2 Contextualised grammar teaching**

Thus, there are, to date, no large-scale studies which investigate the benefits or otherwise of teaching grammar in the context of writing lessons, in which connections are forged for the student writer between the grammar under focus and the learning focus for the writing. However, Hudson (2001) draws attention to a Finnish doctoral study (Laurinen 1955) which reports improved punctuation scores for primary students who have been taught clause structures. Hudson argues that the benefits accrued are because the particular area of grammar taught correlates with the learning focus for

writing, punctuation. Effective punctuation is underpinned by grammatical understanding and the teaching helped the students to make connections between the two.

This synergistic relationship between writing and grammatical understanding is also evident in Fogel and Ehri's (2000) study. This is unusual in taking as its starting point an identified writing problem, the tendency of some ethnic minority children to use non-standard Black English Vernacular (BEV) in their writing. The study set out to "examine how to structure dialect instruction so that it is effective in teaching SE forms to students who use BEV in their writing" (Fogel & Ehri 2000, 215) and found a significant improvement in avoidance of BEV in the group who were given both strategies and guided support. They argue that their results demonstrated that the approach used had "clarified for students the link between features in their own nonstandard writing and features in SE" (2000, 231). The Fogel and Ehri study moves the field forward by beginning to look at the pedagogical conditions which support or hinder the transfer of grammatical knowledge into written outputs. Significantly, too, their study begins with a specific linguistic learning need around which teaching is designed. Fearn and Farnan (2007) have also investigated teaching grammar in the context of teaching writing, seeking to examine if there is "a way to teach grammatical structures that will satisfy high-stakes tests and teachers' needs, and at the same time, positively affect writing performance?" (Fearn & Farnan 2007, 2). Their experimental study encouraged a problem-solving approach, using oral language, and appears to be focused on the use of particular linguistic structures or word class: There is no evident attempt to talk about the construction of meaning or effect through form. Nonetheless, their study did find positive impact of their intervention and their conclusion is that it is beneficial for learners when "grammar and writing share one instructional context" (Fearn & Farnan 2007, 16). A recent meta-analysis by Graham and Perin (2007) looking at effective strategies to teach writing did find that teaching sentence-combining, helping students to construct more structurally complex sentences, had a positive effect. In general, however, there is a dearth of studies which address contextualised grammar teaching probably because "grammar has traditionally been taught and learned in an environment that is devoid of context" (Mulder 2010, 73) or not taught at all.

### **1.3 Theorising grammar-writing connections**

As noted above, a limitation in much of the existing research on grammar teaching is that there is no clear conceptualisation of a theoretical rationale for why grammar might support writing development (indeed much of the research is framed by polemic and ideology). Educational linguists (Carter 1990; Denham & Lobeck 2005; Hancock 2009) contend that a better understanding of how language works in a variety of contexts supports learning in literacy. They draw particularly on the principles of contemporary linguistic theories which are descriptive and socio-cultural in

emphasis, or as Carter describes them, “functionally oriented, related to the study of texts and responsive to social purposes” (Carter 1990, 104). This is in contrast to the more prescriptive approach to grammar which traditional grammars espoused (Hudson 2004). In the US, there has been some emphasis on the notion of grammar in context (Weaver 2006, for example), but a theoretical relationship between grammar and writing has never been adequately articulated, and the idea of ‘in context’ is problematic, often meaning in practice an isolated ‘mini-grammar lesson’ within an English lesson (for a critique of this, see Myhill 2010a).

The difference between prescriptive and descriptive views of grammar is central to a consideration of a theoretical rationale for attention to grammar in the teaching of writing. Prescriptive grammar sets out how language should be used, the rules of language use; whilst descriptive grammar looks at language in use. Denham and Lobeck (2010, 3) contrast linguists who “have sought to build a grammar that would be adequate for describing the language’ with English teachers who ‘have sought to apply a grammar that is already constructed”. Public and political views of grammar tend strongly towards the prescriptive view, maintaining that the role of the teacher is to address grammatical accuracy in writing and eradicate error (Myhill & Jones 2011; Myhill 2011). Hancock, reflecting on the US educational context, observed that “grammar is error and error is grammar in the public mind” (Hancock 2009, 175). In England, the same tendency at public and policy level is evident – grammar is frequently presented as a remediation tool, a language corrective. The Queen’s English Society, whose remit is the preservation of the English language, maintain that grammar is important for the “diagnosing of faults or problems in one’s own writing and in that of others” (QES 2011). Traditional school grammar is prescriptive and is critiqued by Hudson for having “no roots in modern linguistics” and for being “fragmentary, dogmatic and prescriptive” (Hudson 2004, 116). Descriptive theories of grammar counterpoint the normative emphasis on correctness, characteristic of prescriptive grammar, with a more socially-oriented analysis of how language is used, including in different social, linguistic and cultural contexts. A prescriptivist theory of a grammar-writing relationship would argue for the importance of grammar in securing correctness in written expression; a descriptivist theory of a grammar writing relationship would argue for the importance of grammar in illuminating how written text generates meaning in different contexts.

The theoretical approach adopted in this study builds on descriptivist views of grammar. Understanding and analysing how language works in different purposes and contexts makes connections for learners between language as an object of study and language in use, as realised in the act of writing. This is, in effect, a theory of grammar centred upon rhetorical understanding. As a theoretical perspective, this has at its heart the discussion and analysis of how meaning is crafted and created through shaping language to achieve the writer’s rhetorical intentions (Kolln 2002; Locke 2005;

Micciche 2004; Paraskevas 2006). It aims to foster explicit understanding and “conscious control and conscious choice over language which enables both to **see through** language in a systematic way and to use language more discriminatingly” (Carter 1990,119).

A theorised view of grammar teaching in the context of writing which builds on the understandings outlined above, and which focuses on the teaching of writing rather than the teaching of grammar, incorporates the following principles (Myhill 2010a). Firstly, *writing is a communicative act* supporting writers in understanding the social purposes and audiences of texts and how language creates meanings and effects; secondly, *grammar is a meaning-making resource*: supporting writers in making appropriate linguistic choices which help them to shape and craft text to satisfy their rhetorical intentions; and finally, *connectivity*, supporting writers in making connections between their various language experiences as readers, writers and speakers, and in making connections between what they write and how they write it.

#### **1.4 Teachers’ grammatical subject knowledge**

The absence of explicit grammar teaching in the English curriculum in Anglophone countries for nearly 50 years has resulted in many present English teachers not having the grammatical subject knowledge needed to teach grammar confidently. A survey of teachers in England in 1998 (QCA 1998) revealed considerable lack of confidence, in particular with clause structures and syntax. The report noted a “significant gap... in teachers’ knowledge and confidence in sentence grammar and this has implications for... the teaching of language and style in texts and pupils’ own writing” (QCA 1998, 35). In the US context, Vavra (1996) observed the gap between modern linguistics and the prescriptive, rule-bound grammar taught by most English teachers. Cameron (1997) argued that the literature degree qualifications of most English teachers not only leaves them ill-equipped to cope with grammar teaching, but also generates anxiety, hostility, and lack of confidence towards grammar.

This lack of confidence plays out in English classrooms through inaccurate teaching of grammar points (Myhill 2000; Myhill 2003) and insecurity in dealing with students’ questions (Burgess, Turvey, & Quarshie 2000). In two studies investigating pre-service teachers’ engagement with grammar, Cajkler and Hislam (2002, 2004) demonstrate how they struggle with grammatical subject knowledge and how to use it appropriately in the classroom. Hudson (2004) argued that without adequate grammatical knowledge teachers cannot make the analysis of texts explicit, nor can they structure the teaching context effectively.

#### **1.5 Research questions**

The over-arching research question that this study set out to investigate was: What impact does contextualised grammar teaching have upon students’ writing and students’ metalinguistic understanding? The qualitative study, not reported

here, provides evidence concerning students' metalinguistic understanding, as well as complementary evidence about the implementation of the intervention. The randomised controlled trial (RCT) provides evidence concerning the impact of the intervention. Consequently, the following hypotheses were formulated: a) that contextualised grammar teaching will be positively related to students' writing performance; b) that the quality of teachers' grammatical subject knowledge will mediate the impact of contextualised grammar instruction.

## **2.0 Methodology**

### **2.1 Participants**

The participants were teachers ( $n = 32$ ) and students ( $n = 855$ ) of English in 32 different mixed comprehensive schools in the South-West and the Midlands regions of England. In each school, a class of Year 8 students, aged 12-13, formed the sample for the randomised controlled trial. In order to avoid selection bias (Cook and Wei 2002), the school sample was secured by using local authority data to compile a numbered list of all mixed comprehensive schools in the South-West and the Midlands. A random number generator was used to determine a rank order, and each school was approached in rank order until the desired sample of 32 was reached.

Baseline data about participants were collected at school, teacher and student level. At school level, data were compiled on national examination performance, school inspection outcomes, ethnic diversity, Special Educational Needs and number of students entitled to free school meals (as a proxy for socio-economic status). At teacher level, data were collected on years of teaching experience, degree subject studied and gender. In addition, teacher participants undertook a test of grammatical subject knowledge (GSK) at the start of the study and the scores were used as further baseline data. Student data comprised gender, whether they were students with English as an Additional Language (EAL) or entitled to free school meals (FSM). Attainment data in English were also collated, drawing on standard national test results in English at age eleven (including both a writing raw score and writing level) and school predictions of English results in national tests at age fourteen.

### **2.2 Design**

The study sought to investigate whether the use of teaching materials which embedded grammar teaching within teaching units for writing improved students' performance in writing. Additionally, the study sought to examine whether teachers' grammatical subject knowledge (GSK) was a factor upon the efficacy or otherwise of the intervention.

The sample was first stratified at teacher level according to their GSK scores to ensure that the two groups were matched, given that GSK is known to be a factor in the teaching of grammar. The classes were then randomly assigned, using a random number generator, to either a comparison or an intervention group. Because it is not possible in a naturalistic educational setting to prevent any teaching of grammar naturally occurring, we have consistently used the term ‘comparison’ rather than ‘control’ group. The study was blind as participant teachers were not told the research focus was grammar; instead they were told the focus was on the teaching of writing (see also Ethical Considerations 2.7). Full details of the intervention and comparison group teaching are provided further below.

Because causal relationships are rarely deterministic, “to different degrees, all causal relationships are context dependent” (Shadish, Cook & Campbell 2002, 5), a complementary qualitative data set was collated alongside the experimental study to provide in-depth understanding of the theoretical, pedagogical and contextual implications of the statistical data. This mixed method approach is important for RCTs conducted in educational contexts. Indeed, Moore, Graham, and Diamond (2003) argue that “to undertake a trial of an educational or social intervention without an embedded qualitative process evaluation would be to treat the intervention as a black box, with no information on how it worked, how it could be improved, or what the crucial components of the intervention were”. Likewise, Shadish et al. (2002, 71) recommend “the addition of qualitative methodologies to experiments” to provide better interpretation and avoid errors in applying research outcomes to practice. In this study, the data comprised lesson observations, teacher and student interviews, and writing samples. This paper reports principally from the statistical analysis of the RCT (for an overview report of the full study, see Myhill, Jones, Lines & Watson 2012).

### **2.3 Procedure**

Three members of the research team (all former secondary English teachers) devised three teaching units on writing, each addressing a different genre: fictional narrative, argument, and poetry. These teaching units were in harmony with the requirements of the National Curriculum for English (DCSF 2007), a statutory instrument, and addressed teaching objectives for Writing as set out in the Framework for Teaching English (DfES 2001), the recommended policy document guiding English teaching. Three learning objectives which specifically addressed linguistic knowledge were common to all three teaching units, giving the opportunity to explore, for example, how sentence variety may fulfil different purposes in the three genres; the remaining objectives were chosen for their relevance to the genre under study. Each teaching unit was designed to take approximately three weeks of timetabled English lessons, and the study period spanned a school year. In this respect, the intervention was wholly aligned with curriculum and teaching norms and designed for implementation in a naturalistic setting. The teaching units adopted many of the pedagogic practices



common in secondary English classrooms, such as the use of text models, group work and discussion, opportunities for planning and drafting, peer assessment. To that extent, they reflected typical practice in the teaching of writing. However, in addition, within each unit, grammar teaching was embedded, making connections between a linguistic feature and its effect in writing. Some of these were genre-specific. For example, the fictional narrative unit looked at how first or third person are used to create different voice or viewpoint. Others were more generically related to improving writing: for example, varying sentence lengths to create textual rhythm (For more detailed explanation of the teaching activities, see Myhill, Lines & Watson 2011, Myhill 2010b in professional journals). There was no focus on grammatical error or accuracy: rather the focus was to help writers to recognise how making grammatical choices could shape their texts for communicative purposes. Table 1 below provides an overview of the learning objectives and written outcomes for each scheme and provides examples of the embedded grammar focus.

*Table 1: an overview of the three teaching units*

<b>Genre</b>	<b>Learning Objectives</b>	<b>Grammar Focus examples (Intervention only)</b>	<b>Written Outcomes</b>
<b>FICTIONAL NARRATIVE</b>	<ul style="list-style-type: none"> <li>• Varying sentences and punctuation for clarity and effect</li> <li>• Developing varied linguistic and literary techniques</li> <li>• Improving vocabulary for precision and effect</li> <li>• Developing viewpoint, voice and ideas</li> <li>• Using grammar accurately and appropriately</li> </ul>	using sentence fragments or short sentences for emphasis	A story plan, plus writing the opening, the climax, or the resolution of the story.
		use of first/third person to establish viewpoint	
		expanding noun phrases to create character description	
<b>ARGUMENT</b>	<ul style="list-style-type: none"> <li>• Varying sentences and punctuation for clarity and effect</li> <li>• Developing varied linguistic and literary devices</li> <li>• Improving vocabulary for precision and effect</li> <li>• Developing viewpoint, voice and ideas</li> <li>• Structuring, organising and presenting texts in a variety of forms on paper and on screen</li> <li>• Using grammar accurately and appropriately</li> </ul>	using modal verbs to create different degrees of assertion	A written persuasive speech
		using co-ordination to create counter arguments	
		using subordination to articulate propositions	
<b>POETRY</b>	<ul style="list-style-type: none"> <li>• Varying sentences and punctuation for clarity and effect</li> <li>• Developing varied linguistics and literary techniques</li> <li>• Improving vocabulary for precision and impact</li> <li>• Generating ideas, planning and drafting</li> </ul>	using expanded noun phrases to create a picture poem	A portfolio of three poems, plus reflective annotation
		comparing line length in poetry with sentence variety in prose	
		using sentence patterning for effect	

Both the intervention and comparison group were taught the three types of writing over the same period, addressing the same curriculum teaching objectives, and producing the same written outcomes (in other words columns 1, 2, and 4 of

Table 1 were the same for both groups). Both groups were given the same set of stimulus materials and resources, but only the intervention group had detailed teaching units, planned at lesson level, in which grammar was explicitly taught. Thus it is reasonable, as far as is possible within a naturalistic context, to conclude that any differences in writing performance are attributable to the intervention.

#### **2.4 Testing materials**

The impact of the intervention on students' writing performance was determined by a pre and post test sample of writing. Both the pre and post test writing sample were first person narratives, drawing on personal experience, and written under controlled conditions (see Appendix A). In order to avoid any possible bias created by the precise choice of writing task, the topic was selected to avoid known gender preferences in writing and to avoid the need of any specific topic knowledge or experience. To minimise any possible test effect, a cross-over design was used: half the sample took Writing Task 1 as the pretest, and the other half took Writing Test 2; this was reversed for the post-test. Both tests were independently marked by Cambridge Assessment, an independent organisation who were responsible for setting and marking the national test of writing for fourteen year olds until 2006. The mark schemes, one for each task, were developed based on the generic framework underpinning the mark schemes used for all national curriculum writing tests. The model was based specifically on the mark scheme used for the longer writing task in the former Key Stage 3 English test. This criterion-based mark scheme has three strands, with a maximum score of 30 marks: sentence structure and punctuation (max. 8 marks), text structure and organisation (max. 8 marks), and composition and effect (max.14 marks). Each strand is linked to two of the assessment focuses for writing and the criteria are derived from national curriculum levels. For each set of scripts, Cambridge Assessment provided a first marker's set of marks, a second marker's set and a 'resolution mark', adjudicated by a third senior marker if the first two marks were very different. Cambridge Assessment devised training materials for marking; delivered a training day for each marking round; and ensured the usual standardisation checks during the marking. The markers did not know from which treatment group the writing was derived, and the research team was not involved in the process of marking and assessment at all. A full report was provided by Cambridge Assessment on the marking process, including a brief commentary on the writing itself. Markers reported some evidence of carelessness in the post-test writing, which led to speculation that some students might have been less motivated than in the first round, perhaps because the task was completed at the end of the academic year.

#### **2.5 Attrition and Fidelity**

The principle of intervention fidelity is highly problematic in naturalistic educational settings as it is very difficult to control the key variable of the teacher. There is always a methodological trade-off between internal validity and ecological validity, and the study privileged ecological validity. It is important to understand the ways in which the intervention was mediated by the teachers. Thus, although each teacher in the intervention had the same training and the same set of materials, it was neither possible nor ethical to attempt to achieve identical implementation. Teachers were allowed to adapt materials to suit the needs of their students, but they were asked to maintain the specified learning focus and the intervention group were asked to remain as close as possible to the teaching methods in the teaching units. The qualitative study, particularly the lesson observations and teacher interviews, indicates that all but one teacher used the teaching units throughout the research period, and there was a high level of commitment to the project. In all the lessons observed, the teaching units were being used. In the teacher interviews just over half of the teachers specifically claimed they had maintained high fidelity to the teaching schemes, in part because of their commitment to the research project, or as one teacher put it, “I’ve tried to be faithful to your project”. They also appeared content to maintain high fidelity to the intervention because of their approval of the quality of the schemes – “I think the scheme of work is really, really good”; and the sense that they were being successful with the students – “I think the lessons are going really well and I think they’re really learning”. However, there were differences in the way they were implemented. Where adaptations were made, a significant number were pragmatic adaptations in timing, where tasks or activities had to be truncated because of time limitations. However, of particular relevance to the research focus of the study, is the fact that some of the adaptations were altering or omitting the focus on grammar at the heart of the research because it was felt to be too difficult. One teacher reflected that she “found it really difficult and I adapted it to something else”, whilst another “just totally left out the clause part of that exercise, the final element because actually, in the scheme so far we haven’t taught, I haven’t taught them anything about clauses, so it would have just completely have gone over their head”. Another felt that her students “would really struggle with ellipsis. I deliberately didn’t use the term, as with a lot of terms in this scheme, saying a term that they’ve not come across before would scare them”. These adaptations to the grammar element of the lessons link to the statistical finding, reported later in this paper, that teachers’ grammatical subject knowledge was an important factor in mediating the intervention.

However, one teacher in the comparison group showed such low levels of fidelity that it was decided to exclude this class from the data analysis (for example, she taught lessons which did not address writing, and had other teachers teaching her lessons). Thus the final sample for data analysis was 31 teachers and their classes. The initial student sample was 855: over the year of the study, due to students leaving schools or being absent for the post-test, and due to the loss of one whole class, as outlined above, the final student sample was 744.

## 2.6 Analyses

As is common in educational studies, the data arising from this RCT are conceptually hierarchical, comprising individual students grouped within schools/teachers/classes. However, in this case, multilevel modelling is not an appropriate or feasible statistical approach because the intervention or non-intervention was restricted to only a single teacher and their class in each school and applied to the whole of that class. Accordingly, the data were analysed using a multiple regression framework at the single level of the  $n = 744$  individual students, with intervention, school and teacher covariate values attributed to these individuals according to school or class membership.

The response or ‘outcome’ variable considered in the multiple regression modelling was the difference between the post and the pre-test scores (each marked as a percentage of 30) for each participant ( $n = 744$ ). The explanatory covariates comprised the indicator of intervention or non-intervention and then the full set of school, teacher and student baseline measures discussed in Section 2.1. Interaction terms between these covariates were also considered. It should be noted that since these covariates involve a mixture of continuous measures (e.g., teacher grammatical subject knowledge score) and categorical factors (e.g., English as an additional language), one could more specifically refer to the multiple regression modelling as a multi-way analysis of covariance.

A range of stepwise procedures was then used to determine which variables and interactions were statistically significant in impacting the response from amongst the intervention indicator and the other covariates. Final model selection was then checked using full residual analyses and standard diagnostic procedures.

The final results from the model selection indicated that the most appropriate reduced model involved the significant explanatory variables/factors or interactions listed in Table 2 when judged at the 5% significance level ( $p < 0.05$ ). Results relating to this final selected model are reported and discussed in detail in Section 3.

*Table 2: showing most significant covariates at individual, teacher and school level*

Student level	Intervention or comparison group Writing attainment level at age eleven Interaction between writing attainment and intervention	factor: ‘comparison’, or ‘intervention’ factor: ‘below average’, or ‘above average’ factor: ‘comparison & below average’, ‘comparison & above average’, ‘intervention & below average’ or ‘intervention & above average’
Teacher level	Years of teaching experience Linguistic Subject Knowledge	factor: ‘< 5 yrs’, ‘5-10 yrs’, or ‘> 10 yrs’ variable: pre study GSK test score
School level	Latest inspection result Special Educational Needs	factor: ‘poor’, ‘satisfactory’, ‘good or excellent’ variable: % of SEN pupils in school

## 2.7 Ethical Considerations

Blind randomisation, where participants do not know the purpose of the research, requires careful ethical consideration, as it violates the principle of voluntary informed consent. The British Education Research Association ethical guidelines (BERA 2004) recommend that all researchers should “avoid deception or subterfuge unless their research design specifically requires it”. In this case, and particularly because of the contested nature of grammar teaching, knowledge of the grammar focus could have been a significant bias. The BERA guidelines argue that “decisions to use deception or subterfuge in research must be the subject of full deliberation and subsequent disclosure in reporting” (BERA 2004, 6). Prior to conducting the study, institutional ethical approval was sought and secured. At the outset of the study, when participants were invited to give informed consent, they were told that the precise research focus was being withheld from them, and assured that they would be informed at the end of the data collection phase, at which point they could withdraw their consent. At an end of project conference, teachers were given the full details of the purpose of the study, alongside the preliminary findings: in addition all teachers in the comparison group were given the detailed teaching units used in the intervention. When data analysis was complete, all participating teachers were given a summary of the findings of the study and recommendations for classroom practice.

## 3.0 Results

As discussed in Section 2.6, the outcome variable considered was the difference between the post and the pre-test percentage marks for each participant. The mean value of this variable over all 744 participants was 9.24%, indicating an overall improvement between pre- and post- writing tests in both the comparison and intervention groups. In the comparison group ( $n = 332$ ) the mean outcome was 6.41%, lower than in the intervention group ( $n = 412$ ) where mean outcome was 11.52%. So overall, and ignoring the effect of any other covariates, a simple two-sample t-test suggests a highly significant ( $p < 0.001$ ) positive effect size of 5.11 percentage marks for the intervention in terms of improvement in writing attainment. In other words, the embedded teaching of grammar relevant to the writing being studied had an overall beneficial effect on students’ achievement in writing.

However, the simple overall comparison above masks the more complex picture which emerges when significant school, teacher and student explanatory covariates are allowed for via the multiple regression modelling approach discussed in Section 2.6. Table 2 presents coefficient estimates (standard errors, t-values and p-values) relating to the student, teacher and school level variable/factors or interactions for the final model selected from the stepwise regression modelling as described in Section 2.6. All the non-intercept terms in this model are very highly statistically significant (at a 0.1% level) with the exception of the teacher GSK score which nevertheless is still significant at the 5%

level. Factor levels, interactions or variables other than those reported in Table 3 were insignificant at the 5% level. Overall, the model explains only 9% of the raw variation in the pre-and post test differences ( $R^2 = 0.09$ , adjusted  $R^2 = 0.08$ ), but low overall explanatory value is not atypical in educational studies and taken overall this model is very highly significant ( $F_{7,36}$  statistic = 10.42,  $p$ -value = .00001).

Table 3: showing coefficient estimates (standard errors,  $t$ -values and  $p$ -values) at student, teacher and school level for final selected model

<b>Model coefficient</b>	<b>Estimate</b>	<b>Std. error</b>	<b>t-value</b>	<b>p-value</b>
Intercept	-6.1727	(3.5285)	(-1.739)	(<0.100)
<b>Student level</b>				
Above average writing level	-7.4006	(1.9037)	(-3.887)	(<0.001)
Intervention group and above average writing level	8.1246	(1.3863)	(5.860)	(<0.001)
<b>Teacher level</b>				
Teacher Linguistic Subject Knowledge (GSK) score	0.5817	(0.2760)	(2.108)	(<0.050)
Teacher has 5-10 years experience	5.9832	(1.3091)	(4.570)	(<0.001)
<b>School level</b>				
‘Satisfactory’ result in last School inspection	7.0165	(1.9258)	(3.643)	(<0.001)
‘Good or excellent’ in last School inspection	8.3394	(2.0151)	(4.138)	(<0.001)
% Special Educational Needs (SEN) in School	0.6411	(0.1610)	(3.983)	(<0.001)

The interpretation of the ‘intercept’ in this model is that it is the estimated post minus pre-test difference for a ‘baseline’ student who has below average writing ability (whether in the comparison group or the control group), who has a teacher with a low GSK score and with teaching experience of either less than 5 years or of more than 10 years, and who is in a school which received a ‘poor’ result at last inspection and which has a low percentage of students requiring special education needs. Coefficient values other than the intercept in Table 2 then reflect the estimated adjustments to this baseline which apply for students of other types and which are statistically significant different from zero. These effects are discussed in more detail in the following subsections.

Analysis of the raw scores provided by the markers, broken down into the sub-components of the mark schemes, indicates that the greatest improvement was made in Composition and Effect, though similar improvements were made at sentence and text level. This suggests that the impact of the grammar teaching was not simply at the syntactical level of the sentence but had an effect on overall effectiveness, consistent with the pedagogical goals of the intervention.

Table 4: showing improvement by sub-component in the mark schemes

		Total Score	Sentences Structure and Punctuation	Text Structure and Organisation	Composition and Effect
Comparison	Raw scores	6565	1799	1748	3018
	Improvement	806	228	241	337
	% improvement	12.3	12.7	13.8	11.2
Intervention	Raw scores	6446	1764	1704	2978
	Improvement	1256	319	334	604
	% improvement	19.5	18.1	19.6	20.3

### 3.1 Student level effects

The data analysis indicates that the intervention benefitted some writers more than others, and that prior writing attainment was a significant factor in this. For the analysis, students were divided into two attainment groups, ‘more able writers’ and ‘less able writers’ based on prior writing attainment. This was done using the raw score results for Writing in the Key Stage 2 tests, the most fine-grained data we collected was available for all children and which had been externally assessed. The analysis reported in Table 3 shows a significant positive effect on writing performance for students in the intervention group whose baseline data indicated that they were above average in writing attainment. The estimated effect for this group is strong in an educational intervention context (Cohen 1969, 23) and provides robust evidence for the beneficial impact on this group of students of receiving the contextualised grammar teaching. However, the analysis also indicates that it is only those students who were above average in writing at baseline who benefitted strongly from the intervention. The impact of the intervention was neutral on learners whose baseline writing attainment was below average. Table 3 also indicates that more able writers in the comparison group made significantly less improvement in their writing scores over the year than less able writers, suggesting that the teaching of writing may be insufficiently challenging to secure improvement in writers who are already fairly successful. Overall then, more able writers in the intervention group improved significantly more than less able writers who received no significant benefit from the intervention, and more able writers in the comparison group made significantly less improvement over the year than less able writers in either the comparison or intervention groups. This strongly suggests that the intervention may have been particularly well-matched to the learning needs of more able writers.

### 3.2 Teacher level effects

The length of the teacher’s experience and the quality of the teacher’s subject knowledge of grammar were both significant in influencing student writing outcomes, regardless of whether students were in the comparison or the

intervention group. The data suggest that where the intervention was beneficial to students it was most effective with teachers who had between 5 and 10 years teaching experience (ie neither inexperienced nor highly experienced). This may seem paradoxical but it may be that inexperienced teachers lacked the confidence to use the intervention teaching materials appropriately, whilst highly experienced teachers had more deeply embedded pedagogical practices which they found harder to alter, despite the intervention. Certainly, the qualitative evidence from the classroom observations and interviews with teachers suggests that experienced teachers, in particular, made more adaptations to the teaching materials than others, often substituting in their own preferred ways of working. Conversely, the least experienced teachers often followed the teaching materials more rigidly and inflexibly.

Similarly, there is a relationship between grammatical subject knowledge (GSK) and the impact of the intervention. Students in classes with teachers with lower GSK made less improvement than those with teachers with higher GSK. This is a more predictable finding, as the teaching materials required confident mastery of grammar, and although they aimed to provide good support for teachers, it is possible that the grammar knowledge required by the intervention teaching materials was at too high a level for those with lower GSK to teach effectively. Indeed, the lesson observations repeatedly show teachers struggling to cope with student questions on grammar and sometimes communicating incorrect information to students.

### **3.3 School level effects**

At school level, the baseline data collected aimed to capture both social diversity (ethnicity; number of students entitled to free school meals) and school performance (national examination results; school inspection outcomes and the percentage of students with Special Educational Needs). The analysis indicates that the school inspection outcomes and the percentage of students with Special Educational Needs (SEN) were significant factors. The school inspection system in England grades schools through a rigorous process of analysing school performance data on a range of factors, and through observing lessons and grading the teaching quality, and assessing the quality of school leadership. The data indicate that students in schools graded *Good* or *Excellent* through school inspection benefitted more from the intervention than those in schools graded *Satisfactory*. It is important to note that high inspection grades do not correlate with 'high ability' students: all the schools in the study were comprehensive schools, which do not select students on ability, and the inspection assessments take into account the baseline performance of students on entry to the school. So a high-performing school could be achieving good inspection results with lower ability students. It is not surprising that the intervention was more successful in high-performing schools as teaching quality is generally lower in lower-performing schools, which is likely to influence how the intervention materials were used. It is less easy



to account for the influence of SEN, particularly as the data indicate that it was schools with higher percentages of students with SEN which related positively to the beneficial impact of the intervention.

#### **4.0 Discussion**

The study set out to investigate whether contextualised grammar teaching impacts beneficially upon students' writing, and addressed two hypotheses: a) that contextualised grammar teaching will be positively related to students' writing performance; b) that the quality of teachers' grammatical subject knowledge will mediate the impact of contextualised grammar instruction. The analysis indicates that both hypotheses are confirmed. However, the use of multiple regression modelling reveals that the intervention was differentially experienced by learners, with more able writers benefiting strongly whilst weaker writers benefited less.

The lack of impact of the intervention upon less able writers may be because the grammatical features addressed in the teaching materials were poorly matched to their learning needs as writers. One drawback of the RCT was the requirement for standard teaching materials to be used across all groups, making it difficult to tailor the teaching materials to particular classes needs or interest. The design of the intervention materials in this study appear to have been better matched to the learning needs of able writers. It is also possible that more able writers have clearer communicative and rhetorical intentions for their writing than less able writers, enabling them to make more appropriate use of their grammatical understanding to shape text appropriately. Further research could usefully begin by analysing writers' developmental needs, and then designing the teaching schemes around those grammar features most relevant to these needs. It might also be the case that the level of conceptual thinking required to understand grammatical concepts and transfer that learning into their writing was too high a cognitive challenge or created too high a cognitive load for this group. Cognitive load theory (Sweller 1988) acknowledges the constraints that can be placed upon working memory capacity when tackling cognitively complex tasks, such as writing, and argues that teaching materials need to be designed to manage cognitive load appropriately. Similarly, Stark, Kopp and Fischer (2011) argue that extraneous load is "attributed to sub-optimal instructional design and interferes with productive learning". Certainly, the classroom observations suggest that for some students the use of metalinguistic terminology may have been a barrier, rather than a support. This study offers a fruitful impetus for further research which explores the issue of transfer of metalinguistic understanding into writing produced, and also whether appropriate metalinguistic awareness could be generated without the use of metalanguage (for example, by the use of patterns and examples to explain).

The multiple regression modelling also allows us to use the data to construct scenarios which predict outcome effects. A fuller set of these is provided in Appendix B; however, table 5 illustrates some of the key predictive patterns based on the significant covariates. The intervention was of little benefit to less able writers so they are excluded from this scenario-building. For more able writers the possible outcomes are more influenced by the covariates. For example, a student with above average writing ability in a low-performing school with a teacher with low subject knowledge and more than ten years or less than five years' experience is less likely to benefit from the intervention, with writing outcomes improving from -0.6 to 2.1 (or taking the model uncertainty into account, from somewhere in the range -10.7 to -1.3 to somewhere in the range -1.7 to 5.9).

*Table 5: showing predictive scenarios drawing on the multiple regression modelling*

Scenario description	Estimated post minus pre test difference	95% confidence interval
Comparison, high WL, low GSK, <5 or >10 yrs expr, poor LIR, avg SEN	-6.0	(-10.7,-1.3)
Intervention, high WL, low GSK, <5 or >10 yrs expr, poor LIR, avg SEN	2.1	(-1.7, 5.9)
Comparison, high WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	-0.1	(-5.0, 4.8)
Intervention, high WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	8.1	(4.0, 12.2)
Comparison, high WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	11.6	(8.4, 14.8)
Intervention, high WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	19.7	(15.9, 23.5)
<i>WL – Writing Level; GSK – Linguistic Subject Knowledge; LIR – Latest Inspection Result; SEN – Special Educational Needs</i>		

As anticipated in the research design, the results and the predictive scenarios indicate the complexity of interventions in educational settings, which defy simplistic causal explanations. An inherent danger in the use of an RCT is its tendency to be regarded as the 'gold standard' of research (Slavin 2002; Torgerson, Brooks, Porthouse, Burton, Robinson, Wright & Watt 2004), drawing on a medical model of research, and the widespread view that it can answer the question of 'what works' in education. But, as Biesta points out, "a student is not an illness just as teaching is not a cure" (2007, 8) and the closed, deterministic system of the RCT can fail to acknowledge that cause and effect in education are not linear and continuous but open, fluid and dynamic (Morrison 2001). An RCT can provide important evidence and new understandings about a given research 'problem' but it cannot provide evidence which translates directly into classroom practice or policy. Caution is needed in implying "a causal model of professional action" (Biesta 2007, 7) which suggests that mandated action will be implemented in practice and will result in similar

outcomes for all. Understanding “how teachers take on research, adapt it, and make it their own” (Black & William 2003, 633), is both more important and more challenging than simplistic notions of ‘what works’.

In the light of this, the results of this study are important not only in providing robust empirical evidence of the potential benefit of teaching grammar in the context of writing, but also in highlighting the complex interaction effects between student, teacher, school and intervention. This is important in informing appropriate recommendations for policy or professional practice. All interventions in educational settings are always mediated by teacher and school effects, and one-size-fits-all policy imperatives are rarely an appropriate solution to the challenges of raising student attainment. Whilst this study demonstrates the benefit of contextualised grammar teaching on more able writers, it also signals that policy and practice need to consider the level of teachers’ GSK, their experience, different learners’ needs, and the school context in which they operate. Moreover, it opens the way for further experimental and qualitative research which will investigate the pedagogical conditions which support or hinder the transfer of grammatical knowledge into written outputs.

#### **4.0 Conclusions**

Theoretically, this study is important in reframing research in the relationship between grammar teaching and writing. Contrary to previous studies (Andrews et al. 2006; Hillocks 1986, this study offers robust evidence of a positive relationship between grammar and writing. Unlike previous studies, the intervention was predicated upon a theorised pedagogical model for the role of grammar in the teaching of writing which conceptualises grammar as a meaning-making resource and which emphasises connectivity in instructional designs (Myhill 2010a). The study demonstrates that more able students’ learning about writing is enhanced by explicit understanding of how grammar choices can be used to shape written text to satisfy writers’ rhetorical goals, but importantly it also indicated no parallel benefit for less able writers. The results of the intervention are important in indicating future directions for research at the grammar-writing interface.

#### **Acknowledgements**

We are grateful to the Economic and Social Research Council (UK) for the funding which enabled this research to be undertaken.

## References

- Andrews, R., Torgerson, C., Beverton, S., Freeman, A., Locke, T., Low, G., Robinson, A., & Zhu, D. (2006). The effect of grammar teaching on writing development. *British Educational Research Journal*, 32(1), 39-55.
- Bateman, D. R., & Zidonis, F.J. (1966). *The effect of a study of transformational grammar on the writing of ninth and tenth graders*. Champagne, Illinois: National Council of Teachers of English.
- Biesta, G. J. J. (2007). Why 'what works' won't work. Evidence-based practice and the democratic deficit of educational research. *Educational Theory*, 57(1) 1-22.
- Black, P. and Wiliam, D. (2003). 'In Praise of Educational Research': formative assessment. *British Educational Research Journal*, 29(5), 623-37.
- Braddock, R., Lloyd-Jones, R., & Schoer, L. (1963). *Research in written composition*. Urbana: National Council of Teachers of English.
- British Educational Research Association (BERA) (2004). *Revised ethical guidelines for educational research: 2004* Nottingham, UK: BERA.
- Burgess, T., Turvey, A., & Quarshie, R. (2000). Teaching grammar: working with student teachers. *Changing English*, 7(1) 7-21.
- Cajkler, W., & Hislam, J. (2002). Trainee teachers' grammatical knowledge: the tension between public expectations and individual competence. *Language Awareness*, 11(3) 161-177.
- Cajkler, W., & Hislam, J. (2004). Teaching and understanding grammar: how trainee teachers make sense of a new curriculum. In N. Bartels (Ed.) *Applied Linguistics in Language Teacher Education* (pp. 295 -312). Dordrecht: Kluwer Academic Publishers.
- Cameron, D. (1997). Sparing the rod: what teachers need to know about grammar. *Changing English*, 4(2) 229-239.
- Carter, R. (ed.) (1990). *Knowledge about language*. London: Hodder and Stoughton.
- Cohen, J. (1969). *Statistical power analysis for the Behavioral Sciences*. New York: Academic Press.
- Cook, R., & Wei, W. (2002). Selection effects in randomized trials with count data. *Statistics in Medicine*, 21(4) 515-531.
- Denham, K., & Lobeck, A. (eds.) (2010). *Linguistics at school: language awareness in Primary and Secondary Education*. Cambridge: Cambridge University Press.
- Denham, K., & Lobeck, A. (2005). *Language in the schools: integrating linguistics into K12 teaching*. Hillsdale, NJ: Lawrence Erlbaum Associates
- Department for Children, Schools, and Families (DCSF) (2007). *English in the National Curriculum*. <http://curriculum.qcda.gov.uk/key-stages-3-and-4/subjects/key-stage-4/english/index.aspx> . Accessed 10.02.10.
- Department for Education (DfE) (2011) *Interim Results for Key Stage 2 and 3 National Curriculum Assessments in England, 2010/11*. London: DfE. <http://www.education.gov.uk/researchandstatistics/statistics/allstatistics/a00196847/interimresults>. Accessed 20.08.12
- Department for Education and Skills (DfES) (2001). *Framework for teaching English: Years 7, 8 and 9*. London: DfES.
- Dixon, J. (1975). *Growth through English*. Sheffield: NATE.
- Elbow, P. (1981). *Writing with power: techniques for mastering the writing process*. New York: Oxford.
- Elley, W. B., Barham, I. H., Lamb, H., & Wylie, M. (1979). *The role of grammar in a secondary school curriculum: Educational Research Series No 60*. Wellington: New Zealand Council for Educational Research.
- Elley, W. B., Barham, I. H., Lamb, H., & Wylie, M. (1975). The role of grammar in a secondary school curriculum. *New Zealand Council for Educational Studies*, 10, 26-41.
- EPPI Review Group for English (2004). *The effect of grammar teaching (syntax) in English on 5 to 16 year olds' accuracy and quality in written composition*. London: EPPI-Centre.
- Fearn, L., & Farnan, N. (2007). When is a verb using functional grammar to teach writing. *Journal of Basic Writing*, 26(1) 1-26.
- Fogel, H., & Ehri, L. C. (2000). Teaching elementary students who speak black English vernacular to write in Standard English: effects of dialect transformation practice. *Contemporary Educational Psychology*, 25, 212-235.

- Gordon, E. (2005). Grammar in New Zealand schools: two case studies. *English Teaching: Practice and Critique*, 4(3) 48-68.
- Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools – A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Hancock, C. (2009). How linguistics can inform the teaching of writing. In R. Beard., D. Myhill., J. Riley., & M. Nystrand (Eds), *The Sage handbook of writing development* (pp.194-208). London: Sage.
- Hillocks, G., & Smith, M. (1991) Grammar and usage. In J. Flood., J. M. Jensen., D. Lapp., & J. R. Squire (Eds.), *Handbook of research on teaching the English Language Arts* (pp. 591-603). New York: MacMillan.
- Hillocks, G. (1986). *Research on written composition*. Urbana, IL: ERIC Clearing House on Reading and Communication Skills.
- Hudson, R. (2004). Why Education needs linguistics. *Journal of Linguistics*, 40 (1) 105-130.
- Hudson, R. (2001). Grammar teaching and writing skills: the research evidence. *Syntax in the Schools*, 17, 1-6.
- Keith, G. (1997). Teach yourself English grammar. *The English and Media Magazine* No 36, Summer 1997.
- Kolln, M. (2002). *Rhetorical grammar: grammatical choices, rhetorical effects*. New York: Longman.
- Laurinen, I. (1955). *The development of sentence sense in the light of the results attained in the teaching of writing in Finnish primary schools*. Unpublished PhD thesis.
- Locke, T. (2005). Writing positions and rhetorical spaces. In B. Doecke and G. Parr (Eds.), *Writing = Learning* (pp. 75-95). AATE, Wakefield Press: South Australia..
- Locke, T. (2009). Grammar and writing – the international debate. In R. Beard., D. Myhill., J. Riley., & M. Nystrand (Eds.), *The Sage Handbook of Writing Development* (pp. 182-193). London: Sage.
- Micciche, L. (2004). Making a case for rhetorical grammar. *College Composition and Communication*, 55(4), 716-737.
- Moore, L., Graham, A., & Diamond, I. (2003), On the feasibility of conducting randomized comparison trials in Education. *British Education Research Journal*, 29(5) 673-689.
- Morrison, K. (2001). Randomised controlled trials for evidence-based education: some problems in judging ‘What Works’. *Evaluation and Research in Education*, 15(2) 69-83.
- Mulder, J. (2010). Envisioning linguistics in secondary education: an Australian exemplar . In A. Denham & K. Lobeck (Eds.), *Linguistics at school: language awareness in Primary and Secondary Education*. Cambridge: Cambridge University Press.
- Myhill, D. A. (2010a). Ways of knowing: grammar as a tool for developing writing. In T. Locke (Ed.), *Beyond the grammar wars: A resource for teachers and students on developing language knowledge in the English/literacy classroom* (pp. 129 - 148). London: Routledge
- Myhill, D. A. (2011). Living language, live debates: grammar and Standard English. In J. Davison, C. Daly, & J. Moss (Eds.), *Debates in English Teaching* (pp. 63-77). London: Routledge.
- Myhill, D. A. (2000). Misconceptions and difficulties in the acquisition of meta-linguistic knowledge. *Language and Education*, 14(3) 151-163.
- Myhill, D. A. (2003). Principled understanding? Teaching the active and passive voice. *Language and Education*, 17(5) 355-370.
- Myhill, D. A. (2010b). Harnessing grammar: weaving words and shaping texts. *Better: Evidence-Based Education*, Winter 2010 pp12-13.
- Myhill, D. A., & Jones, S. M. (2011). Policing grammar: the place of grammar in literacy policy. In A. Goodwyn & C. Fuller (Eds.), *The Literacy Game* (pp. 45-62). London: Routledge
- Myhill, D. A., Jones, S. M., Lines, H. E., & 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.
- Myhill, D. A., Lines, H. E., & Watson, A. (2011). Making meaning with grammar: a repertoire of possibilities *mETaphor*. 2:1-10
- National Commission on Writing in America’s Schools and Colleges (NCW) (2003). *The neglected ‘R’: The need for a writing revolution*. New York: College Board.
- OFSTED (2009). *English at the crossroads*. London: OFSTED.

- Paraskevas, C. (2006). Grammar apprenticeship. *English Journal*, 95(5) 65 -69.
- Qualifications and Curriculum Authority (QCA) (1998). *The grammar papers*. London: QCA.
- Queen's English Society (QES) (2011). [www.queens-english-society.com/about.aspx](http://www.queens-english-society.com/about.aspx). Accessed 20.01.11
- Robinson, N. (1959). *The relation between knowledge of English grammar and ability in English compositions*. Unpublished MEd thesis, University of Manchester
- Salahu-Din., D. Persky., H., & Miller, J. (2008). *The nation's report card: writing 2007*. Washington DC: Institute of Education Sciences, US Department of Education.
- Shadish, W.R., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Slavin, R. E. (2002). Evidence-based education policies: transforming educational practice and research. *Educational Researcher*, 31(7) 15-21.
- Stark, R., Kopp, V., & Fischer, M.R.. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. *Learning and Instruction*, 21(1) 22-33.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-285.
- Torgerson, C., Brooks, G., Porthouse, J., Burton, M., Robinson, A., Wright., K., & Watt, I. (2004) *Adult literacy and numeracy interventions and outcomes: a review of controlled trials*. London: National Research and Development Centre for Adult Literacy and Numeracy.
- Van Gelderen, A. (2006). What we know without knowing it: Sense and nonsense in respect of linguistic reflection for students in elementary and secondary education. *English Teaching: Practice and Critique*, 5(1), 44-54.
- Van Gelderen, A., & Oostdam, R. (2005). Effects of fluency training on the application of linguistic operations in writing. *L1 Studies in Language and Literature*, 5(2) 215-240.
- Vavra, E. (1996). On not teaching grammar. *English Journal*, 85(7) 32-37.
- Weaver, C. (1996). Teaching grammar in the context of writing. *English Journal*, 85(7) 15-24.
- Wyse, D. (2004). Grammar. For writing? A critical review of empirical evidence. *British Journal of Educational Studies*, 49(4) 411-427.

## APPENDIX A

### Writing Task 1: Challenges

Challenges come in many shapes and sizes, from adrenaline-pumping adventures like rock climbing or bungee jumping to more everyday events like starting a new school, overcoming a fear, getting into a team or learning a new skill. Often a challenge is small to someone else but very big to you – for example, having to say sorry for something you have done or meeting someone you don't know for the first time.

**Write an account of a challenge you have faced in your life so far for a school magazine feature on 'Challenging Situations'.**

You could:

- Describe the challenge.
- Tell the story of how you tackled it and what happened in the end.
- Reflect on how you feel about this now
- Choose language that will make your account vivid and interesting for someone else to read

You *might* want to start like this:

*At the time, it seemed like the biggest challenge of my life....*

### Writing Task 2: Childhood Fears

All of us get frightened about things from time to time – especially when we are young children. Our fears may be based on something real that has happened or something we only imagine will happen. For example, it could have been a fear of something imaginary, such as serpents under the bed, or a very real fear, such as a fear of flies.

**Write about your childhood fears, real or imagined, for a school magazine feature on 'Things that frighten us when we are small'.**

You could:

- Choose a time when you were frightened or nervous about something
- Describe it in detail
- Explain why you were afraid and what happened in the end
- Reflect on how you feel about this now
- Choose language that will make your account vivid and interesting for someone else to read

You *might* want to start like this:

*I don't think I have ever been as scared as when...*

APPENDIX B

Scenario description	Estimated post minus pre test difference	95% confidence interval
Comp, high WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	-6.0	(-10.7,-1.3)
Comp, high WL, mid GSK, 5-10 yrs expr, poor LIR, avg SEN	-3.7	(-7.9, 0.5)
Comp, high WL, high GSK, 5-10 yrs expr, poor LIR, avg SEN	-1.4	(-6.1, 3.3)
Comp, high WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	-0.1	(-5.0, 4.8)
Comp, high WL, low GSK, <5 (-2.1, 4.1)or >10 yrs expr, good or excel LIR, avg SEN	1.0	(-2.1, 4.1)
Intervent or Comp, low WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	1.4	(-3.3, 6.1)
Intervent, high WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	2.1	(-1.7, 5.9)
Comp, high WL, mid GSK, 5-10 yrs expr, poor LIR, avg SEN	2.3	(-2.0, 6.6)
Comp, high WL, mid GSK, 5-10 yrs expr, good or excel LIR, avg SEN	3.3	( 1.3, 5.3)
Intervent or Comp, low WL, mid GSK, 5-10 yrs expr, poor LIR, avg SEN	3.7	(-0.7, 8.1)
Intervent, high WL, mid GSK, 5-10 yrs expr, poor LIR, avg SEN	4.4	(1.0, 7.8)
Comp, high WL, high GSK, 5-10 yrs expr, poor LIR, avg SEN	4.6	(-0.2, 9.4)
Comp, high WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	5.6	(2.8, 8.4)
Intervent or Comp, low WL, high GSK, 5-10 yrs expr, poor LIR, avg SEN	6.0	(1.0, 11.0)
Intervent, high WL, high GSK, 5-10 yrs expr, poor LIR, avg SEN	6.7	(2.4, 11.0)
Comp, high WL, low GSK, 5-10 yrs expr, good or excel LIR, avg SEN	7.0	(3.6, 10.4)
Intervent or Comp, low WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	7.3	(2.1, 12.5)
Intervent, high WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	8.1	(4.0, 12.2)
Intervent or Comp, low WL, low GSK, 5-10 yrs expr, good or excel LIR, avg SEN	8.4	(4.6, 12.2)
Intervent, high WL, low GSK, 5-10 yrs expr, good or excel LIR, avg SEN	9.1	(6.1, 12.1)
Comp, high WL, mid GSK, 5-10 yrs expr, good or excel LIR, avg SEN	9.3	(6.8, 11.8)
Intervent or Comp, low WL, mid GSK, 5-10 yrs expr, poor LIR, avg SEN	9.7	(4.8, 14.6)
Intervent, high WL, mid GSK, 5-10 yrs expr, poor LIR, avg SEN	10.4	(6.7, 14.1)
Intervent or Comp, low WL, mid GSK, 5-10 yrs expr, good or excel LIR, avg SEN	10.7	(7.5, 13.9)
Intervent, high WL, mid GSK, 5-10 yrs expr, good or excel LIR, avg SEN	11.4	(9.0, 13.8)
Comp, high WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	11.6	(8.4, 14.8)
Intervent or Comp, low WL, high GSK, 5-10 yrs expr, poor LIR, avg SEN	12.0	(6.5, 17.5)
Intervent, high WL, high GSK, 5-10 yrs expr, poor LIR, avg SEN	12.7	(8.2, 17.2)
Intervent or Comp, low WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	13.0	(9.0, 17.0)
Intervent, high WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	13.8	(10.4, 17.2)
Intervent or Comp, low WL, low GSK, 5-10 yrs expr, good or excel LIR, avg SEN	14.3	( 9.8, 18.8)
Intervent, high WL, low GSK, 5-10 yrs expr, good or excel LIR, avg SEN	15.1	(11.7, 18.5)
Intervent or Comp, low WL, mid GSK, 5-10 yrs expr, good or excel LIR, avg SEN	16.7	(12.7, 20.7)
Intervent, high WL, mid GSK, 5-10 yrs expr, good or excel LIR, avg SEN	17.4	(14.5, 20.3)
Intervent or Comp, low WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	19.0	(14.3, 23.7)
Intervent, high WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	19.7	(15.9, 23.5)
<i>WL – Writing Level; GSK – Linguistic Subject Knowledge; LIR – Latest Inspection Result; SEN – Special Educational Needs</i>		



Table 1: an overview of the three teaching units

Genre	Learning Objectives	Grammar Focus examples (Intervention only)	Written Outcomes
<b>FICTIONAL NARRATIVE</b>	<ul style="list-style-type: none"> <li>Varying sentences and punctuation for clarity and effect</li> <li>Developing varied linguistic and literary techniques</li> <li>Improving vocabulary for precision and effect</li> <li>Developing viewpoint, voice and ideas</li> <li>Using grammar accurately and appropriately</li> </ul>	using sentence fragments or short sentences for emphasis	A story plan, plus writing the opening, the climax, or the resolution of the story.
		use of first/third person to establish viewpoint	
		expanding noun phrases to create character description	
<b>ARGUMENT</b>	<ul style="list-style-type: none"> <li>Varying sentences and punctuation for clarity and effect</li> <li>Developing varied linguistic and literary devices</li> <li>Improving vocabulary for precision and effect</li> <li>Developing viewpoint, voice and ideas</li> <li>Structuring, organising and presenting texts in a variety of forms on paper and on screen</li> <li>Using grammar accurately and appropriately</li> </ul>	using modal verbs to create different degrees of assertion	A written persuasive speech
		using co-ordination to create counter arguments	
		using subordination to articulate propositions	
<b>POETRY</b>	<ul style="list-style-type: none"> <li>Varying sentences and punctuation for clarity and effect</li> <li>Developing varied linguistics and literary techniques</li> <li>Improving vocabulary for precision and impact</li> <li>Generating ideas, planning and drafting</li> </ul>	using expanded noun phrases to create a picture poem	A portfolio of three poems, plus reflective annotation
		comparing line length in poetry with sentence variety in prose	
		using sentence patterning for effect	

Table 2: showing most significant covariates at individual, teacher and school level

Student level	Intervention or comparison group Writing attainment level at age eleven Interaction between writing attainment and intervention	factor: 'comparison', or 'intervention' factor: 'below average', or 'above average' factor: 'comparison & below average', 'comparison & above average', 'intervention & below average' or 'intervention & above average'
Teacher level	Years of teaching experience Linguistic Subject Knowledge	factor: '< 5 yrs', '5-10 yrs', or '> 10 yrs' variable: pre study GSK test score
School level	Latest inspection result Special Educational Needs	factor: 'poor', 'satisfactory', 'good or excellent' variable: % of SEN pupils in school

Table 3: showing coefficient estimates (standard errors, t-values and p-values) at student, teacher and school level for final selected model

Model coefficient	Estimate	Std. error	t-value	p-value
Intercept	-6.1727	(3.5285)	(-1.739)	(<0.100)
<b>Student level</b>				
Above average writing level	-7.4006	(1.9037)	(-3.887)	(<0.001)
Intervention group and above average writing level	8.1246	(1.3863)	(5.860)	(<0.001)
<b>Teacher level</b>				
Teacher Linguistic Subject Knowledge (GSK) score	0.5817	(0.2760)	(2.108)	(<0.050)
Teacher has 5-10 years experience	5.9832	(1.3091)	(4.570)	(<0.001)
<b>School level</b>				
'Satisfactory' result in last School inspection	7.0165	(1.9258)	(3.643)	(<0.001)
'Good or excellent' in last School inspection	8.3394	(2.0151)	(4.138)	(<0.001)
% Special Educational Needs (SEN) in School	0.6411	(0.1610)	(3.983)	(<0.001)

Table 4: showing improvement by sub-component in the mark schemes

		Total Score	Sentences Structure and Punctuation	Text Structure and Organisation	Composition and Effect
Comparison	Raw scores	6565	1799	1748	3018
	Improvement	806	228	241	337
	% improvement	12.3	12.7	13.8	11.2
Intervention	Raw scores	6446	1764	1704	2978
	Improvement	1256	319	334	604
	% improvement	19.5	18.1	19.6	20.3

Table 5: showing predictive scenarios drawing on the multiple regression modelling

Scenario description	Estimated post minus pre test difference	95% confidence interval
Comparison, high WL, low GSK, <5 or >10 yrs expr, poor LIR, avg SEN	-6.0	(-10.7,-1.3)
Intervention, high WL, low GSK, <5 or >10 yrs expr, poor LIR, avg SEN	2.1	(-1.7, 5.9)
Comparison, high WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	-0.1	(-5.0, 4.8)
Intervention, high WL, low GSK, 5-10 yrs expr, poor LIR, avg SEN	8.1	(4.0, 12.2)
Comparison, high WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	11.6	(8.4, 14.8)
Intervention, high WL, high GSK, 5-10 yrs expr, good or excel LIR, avg SEN	19.7	(15.9, 23.5)

WL – Writing Level; GSK – Linguistic Subject Knowledge; LIR – Latest Inspection Result; SEN – Special Educational Needs