

Key Stage 3 English Y9 Tests, 2004

Briefing paper for headteachers and governors

September 2004

Although there has been dissatisfaction with the reliability of the KS3 English tests since they were introduced, a number of factors (set out below) have contributed to this year's more than usually flawed testing process. For the first time, the DfES agree: at the time of writing, ministers have not accepted the results from QCA and they are not yet published.

This briefing paper sets out the problems. It gives the headlines for each contributory factor and provides additional detail, drawing on feed-back from heads of English and Devon's team of KS3 English consultants.

It is important that schools understand the factors which make the testing process unreliable, especially when judgements about schools' performance are based on this test and comparisons are made with performance in maths and science. And, given the importance placed on SATs results, it is very hard for pupils and parents to understand why there should be this unreliability within the English tests; they are justifiably bewildered by reports of inconsistent and inaccurate marking. It is hard on schools and pupils who have justifiably performed well in the tests to feel their results are devalued by a flawed process.

Widespread understanding of these issues will enable schools and their governing bodies to make reasoned representations to the DfES and Members of Parliament about the situation faced by schools this year and in arguing for a fairer and more reliable approach to assessment in English .

Separate markers for the reading and writing components

There was general welcome for the introduction for separate markers for the reading and writing elements as it was felt this would lead to fairer marking. However marking is still very varied in quality (see below) and the failure of the electronic system of collecting marks led to late publication of results and papers being returned to schools without the usual borderline checks.

1. The test papers

➤ *Shakespeare paper*

Lack of equivalence of difficulty in the questions with stark discrepancy in the level of challenge: the Macbeth question (*How do these extracts explore the idea that it is difficult to know whom to trust?*) was much more challenging than the questions set for Henry V (*What advice would you give to help the actor playing Henry to convey his different moods before and after the battle?*) and Twelfth Night (*Explain whether you think Malvolio deserves sympathy in these extracts, and why*). The Macbeth question was not a suitable vehicle for enabling 14 year olds to demonstrate their knowledge and understanding of the play and the set scenes. Shakespeare is assessed through coursework at GCSE and this allows a more varied and differentiated approach to be taken.

Questions are trialled with Year 9 pupils but the reliability of this trialling has to be questioned in this case. There is clear evidence that the Macbeth question failed to elicit and judge pupils' understanding of the play through the two set scenes: in a number of schools, pupils who achieved 29 or 30 out of 32 for the reading paper, were getting 0-3 marks out of 18 for their Shakespeare answer; in one centre, a third of the year group received 0 for their Macbeth answer, which is a clear indication of pupils being disadvantaged by the question.

➤ *Longer writing task – pupils were put in the role of a writer of a series of detective novels, given some notes for a first chapter and then asked to write the first chapter*

There was mixed response to this question from heads of English. Some considered the detective novel a specific and rather complex genre for Year 9 pupils. There was more consensus that the design of the question was misleading, with background information and cluttered question layout confusing lower attaining pupils.

➤ *Shorter writing task*

There continues to be a problem with the pairing of the shorter writing task with the question on the Shakespeare play (this will not happen next year when this element will be paired with the longer writing task). The content of the task is loosely linked to each play e.g. a thematic link, but some pupils, understandably, are confused by this. Instead of focusing on crafting their writing, they are distracted by trying to make links to the play.

There was an added complication this year with pupils having to write the answers to the two sections in different answer booklets – a distraction of some pupils.

➤ *Reading paper*

This paper was generally considered fair although several heads of English were surprised that the texts were entirely non-fiction and there was no pre-1914 writing. Teachers were led to believe that there would be both literary and pre-1914 texts on this year's paper. However, there was an unannounced change in weighting in the reading mark scheme with fewer marks for straightforward skills of information retrieval and greater emphasis on the more challenging skills related to the analysis and explanation of writers' effects; schools were not told about this in advance.

2. The quality of marking

Marking of English inevitably has a subjective element. Current mark schemes are very specific and structured. However, if markers are not currently teaching English in the state sector, they may lack the subject knowledge to interpret mark schemes in the light of pupils' responses. It was clear that this was the case this year.

Devon English consultants who supported schools in the appeals process this year, found inconsistencies in marking by the same marker in a single school's scripts, particularly in the writing papers. This included very harsh marking of the composition and effect element in the longer writing task; in other cases, the marker did not reward variety in sentence length and construction when there was clear evidence for it. The writing of high attainers was unevenly rewarded. In other schools, inconsistencies were noted where some scripts were marked over-generously.

Longer writing task – the detective story

Aspects of the mark scheme were closely related to the detective genre. This meant that pupils who demonstrated specific abilities in narrative writing (part of the NC requirements) but did not include all elements of the detective genre (not an NC requirement) did not get appropriately rewarded when the mark scheme was narrowly interpreted. In addition, the examples of pupils' writing, included to illustrate the application of the mark scheme, only exemplified one of the three aspects being marked, sentence structure & punctuation. This added to the tendency of unskilled markers to undervalue text structure & organisation and composition & effect.

Shorter writing task

Again, the examples of pupils' writing only exemplified application of the sentence structure aspect of the mark scheme.

Reading paper

The least problematic area: there was some harsh marking with narrow interpretation of the mark scheme.

3. Administration

The software system (designed by Pearsons) for the electronic collection of marks did not function properly; this meant that papers were returned without borderline checks. The borderline checking involves the marker in running a clerical check and reviewing the fairness of judgements, when the candidate is within 2 marks of a borderline.

The intention was to gather marks electronically from the markers of the reading and writing papers, collate them and identify papers for borderline checks. The failure of the software

meant collection of the marks was delayed and the imperative to meet the deadline for publication of results to schools led to papers being returned without these checks.

In the event, results were published late and this made statutory requirement to report to parents very difficult for schools.

4. Appeals

A number of schools spent an inordinate amount of time and money going through the process of script review (appeal against marking) at the end of the Summer term. In some cases, two people were taken off timetable for a day to complete the scrutiny and paperwork for a review. On return to school in September, they then found that they were invited to send all their scripts back anyway for borderline checking.

It is intolerable that schools have to bear the onus of remedying the inadequacies of another organisation.

5. An alternative to tests

Five Devon schools are currently involved in an assessment project led by QCA and the KS3 National Strategy. The project is developing materials and approaches for reliably assessing pupils' on-going English work at Key Stage 3.

The suggested approaches are designed to support teacher assessment so that

- it is systematic, accurate and linked to national standards
- enables teachers to monitor pupils' progress in order to inform teaching
- enables teachers to find out more about the patterns of pupils' progress in KS3 e.g. why some pupils make little or not progress

A variety of methods of assessment are suggested, supported by optional tasks

The materials are still at the development and trialling stage but the project does offer a practical way forward in addressing the complexities of assessing attainment in reading and writing and in increasing the credibility of teacher assessment.

KS3 EDP Group
Devon LEA
September 2004