

OFSTED Invitation Conference: Assessment in Key Stage 3 History

British Library, March 19th 2002.

At this conference fifty teachers from different LEAs joined HMI and QCA and DfES officers with subject responsibility in order to discuss how to improve assessment of National Curriculum history. The two main sessions began with presentations of the work done by the history departments at Snaith School, East Riding, and Matthew Moss School, Rochdale. Workshop groups then took up the issues raised by these speakers in order to establish some principles for good assessment, and in particular to address two questions:

- *How far are National Curriculum levels useful for within-key stage as well as end-of-key stage assessment?*
- *How can assessment and assessment data be used to raise standards?*

Assessment of history: OFSTED findings

The teaching of history in Key Stage 3 is generally good, amongst the best of all subjects. Assessment in history is good in just over four schools in ten, satisfactory in a little more than that, but unsatisfactory in about one in ten. Areas of relative strength include day-to-day marking and oral feedback to pupils. In general preparation for GCSE examinations is a good feature of the work of departments, but this is not always transferred to aspects of assessment in Key Stage 3.

In a recent survey of schools where assessment in history had been identified as good, the following conclusions were drawn.

- The best departments make significant and effective use of assessment to 'empower pupils with the tools for raising standards'. Assessment activity is integral to their work, including planning, teaching, assessing through a range of modes, and evaluating on the basis of assessment information.
- There is a strong correlation between schools with well-developed policies and established procedures, and departments where assessment is effective, consistent and manageable.
- Where assessment practice is good, pupils generally considered themselves to be well informed about their attainment and progress. Clear learning objectives and detailed marking were important in this respect. Performance grades and targets were seen to be most helpful where they provided sufficiently precise information to enable pupils to address weaknesses or realise next steps in learning.
- Where sufficient time was given in lessons for discussion of assessment criteria or outcomes, this was valuable in focusing pupils' attention on what they need to do next or how to improve.

- Good departments take end of key stage assessment seriously and have established procedures to promote consistency and accuracy.

The aspect of assessment that historians have more difficulty with is the assessment of progress over time, specifically whether and how to use National Curriculum levels within Key Stage 3. More broadly, in a minority of schools insufficient use is made of assessment to raise standards.

Conference Session 1. The use of National Curriculum levels for within key-stage assessment

The issues

OFSTED evidence shows that many schools are now asking for periodic reporting of pupils' progress by National Curriculum levels, either annually or even bi-annually or termly. Sometimes, such requests make assumptions about the transferability of assessment methods from one subject to another which are not borne out in practice. History departments have responded to such requests in diverse ways—most frequently, they have established periodic assessments to provide them with information on pupils' attainment, using tests, assignments, or broader progress reviews.

Many schools (and some LEAs) have sought to re-write the National Curriculum levels to provide finer-tuned and 'pupil-friendly' statements. More often than not, these are in themselves problematic, ascribing to the 'best fit' level descriptions a specificity that they were never intended to imply. To make this picture even more confusing, different schools come up with different interpretations. Some have assumed that National Curriculum levels are of the same kind as those used in a GCSE levels of response mark scheme: although the latter are in fact context- and task-specific, and very fine-tuned, leading to an award of marks which, only when taken as a whole, yield a grade. In some schools, the original concept of a level as representative of two years' intellectual progression has been forgotten, as have the norms for average attainment at the end of the Key Stage and the equivalence of level 6.5 to GCSE Grade C. The attempt to calibrate re-written National Curriculum levels against GCSE grade descriptors is also problematic.

Some assessment tasks are in themselves problematic, often where teachers have attempted to assess too many items, or where they have attempted to differentiate by task or by stepping. In either case, there is a danger that the assignments produce fragmentary rather than holistic results. Where, additionally, teachers have attempted to mark each item with levels, the outcomes are inconsistent and meaningless.

Conference presentation and discussion

In the opening presentation, Steven Elsley, head of history at Snaith School, demonstrated how his department uses a range of assignments across the key stage, each targeted at one or two principal objectives and additionally subsidiary objectives, with each assignment therefore leading to a holistic overview of performance.

Discussion in conference workshops reached the following conclusions about the factors which a department needs to take into account.

- Unlike some National Curriculum subjects that have separate strands and relatively straightforward learning steps, learning and its assessment in history is 'messy'. Attempts to impose a precise discipline on history assessment are likely to be counter-productive.
- The deconstruction of level descriptions (for example, the isolation of particular statements relating to one 'strand') does not produce a set of learning steps that can readily be used as a mark scheme. Mark schemes, in general, need to be task-specific, and can only be calibrated against the levels in the context of other factors, for example the level of difficulty of source material or the quality of deployment of knowledge.
- There is much duplication of activity within schools, which do not feel confident about national guidance to date. More, authoritative guidance and exemplification would be valuable.

The workshops established the following broad principles about effective within-key stage assessment.

- Assignments which bring together learning at the end of a section can produce valid assessments; this might include, for example, the outcome of a structured but independent enquiry into a significant aspect of the study, drawing from across the knowledge, skills and understanding of National Curriculum history. Such assessments support the concept of 'best fit' use of level descriptions.
- Where assignments are targeted at specific objectives, such as causation or evaluation of sources, these should be seen to provide only indicative information about pupils' attainment, and need to be set in the context of broader performance.
- Where teachers have worked together to develop their understanding of 'levelness', this can lead to consistent assessment within departments based principally on good professional judgement, supported by evidence. However, in order to promote accurate and reliable assessment between schools, much more needs to be done to promote standardisation, either via the National Curriculum website or building upon local networks.

Conference Session 2. How can assessment and assessment data be used to raise standards?

The issues

OFSTED evidence suggests that in some schools inadequate use is made of assessment processes and outcomes to raise standards. In these cases, assessment is something 'extra', insufficiently prepared for in the teaching and inadequately exploited so that both pupils and teachers gain from the experience through subsequent review and action. Additionally, some systems are very burdensome, but to little good effect. For example, some very laborious marking fails

to help pupils improve their work, and some negligent marking can reinforce bad habits; sometimes, too, marking varies unacceptably within departments.

Too often, assessment is something that is done to pupils rather than something in which they are involved in every respect. Pupils are unaware of the objectives and criteria against which they are being assessed, or why they have received the grade, mark, level or comment that appears on their work.

The use of data is now an expectation in most schools, but sometimes data collection appears to be an end in itself rather than as information to be interrogated as a means to raise performance.

Conference presentation and workshops

Steven Jolly, Head of History at Matthew Moss School, Rochdale, offered a wide range of strategies used by the department to involve pupils and teachers in the assessment process and its evaluation, so helping to explain the fast improving standards demonstrated by pupils in GCSE history at the school.

Workshop groups reached the following conclusions about the principles of good assessment practice.

- Marking of pupils' work presents the best opportunity for formal, individualised comments that over time build up a picture of progress. The marking load needs to be manageable, so if marking is to be informative it must be selective and focused, principally commenting on learning objectives, both strengths and weaknesses. The importance of measured oral feedback, both in its own right and as a supplement to marking, should not be underestimated.
- Schools use many systems for marking and grading pupils' attainment and effort; these notations used are less important than that they are consistently applied and serve to raise pupils' expectations and performance.
- Assessment should be integral to the teaching and learning that has taken place, with pupils being exercised in the knowledge, skills and understanding of National Curriculum history that they have learned. They should know what the objectives of the assessed work are at the levels at which they can be expected to achieve.
- Level descriptions do not yield helpful objectives or subsequent learning targets for individual pieces of work. Such finer-tuned objectives are best when they are task- and context-specific. So, for example, although one may create a hierarchy of statements describing progress in pupils' understanding of causation in history, in practice the level of demand will vary greatly according to the complexity of the issue. The application of this hierarchy to levels is, therefore, dependant on the outcomes of pupils' work, which are more suitable for calibration against the levels than the application of any prescribed formula.
- It is important that pupils are involved in their work, but such 'self-assessment' needs to be focused. Pupils need to know the objectives of the task and to be given a clear view of how far they have achieved those objectives, and what remains to be improved. Often, these targets will be small steps, although ideally pupils should know how their work corresponds to the demands of the National Curriculum levels.

- Pupils learn in many ways, and assessment needs to recognise this, and especially that an emphasis on only written outcomes disadvantages particular groups of pupils.
- Assessment should not be a device to catch pupils out. Whatever the mode employed, such as a test or assignment, pupils should know what it expected of them. The modelling of good work, and review to evaluate answers, can be important levers in raising pupils' expectations.
- Assessed work provides valuable information about the effectiveness of teaching as well as pupils' attainment. It is a powerful evaluative tool that should influence subsequent teaching and curriculum review, including planning for progression across the key stage.
- Performance data is useful both to inform teaching and as a tool to inform departmental management. Whilst there is no requirement for pupils to be awarded a level in history at the end of Key Stage 2, schools do receive data on pupils' performance which is pertinent to history teachers as an indicator of pupils' overall ability and their attainment in English.
- Performance data is a critical management tool to review the performance of the department as a whole, between classes and in comparison with other curriculum areas. Some departments are supplementing whole school data with more detailed departmental records, including the relative performance of different groups of pupils.
- Assessment is a demanding aspect of teaching, and requires strong leadership and good management if it is to be effective. Departments need a policy that spells out the why and how of assessment, including marking. Consistency is promoted by monitoring and sharing the work of colleagues and by establishing sound benchmarks of standards, perhaps through the compilation of a department portfolio of assessed work, which can be compared more widely. Maintaining procedures for reliable and accurate end of key stage judgements are central to the leadership of a professional department.