

The role of evidence in the new KS4 National Curriculum for England and the AQA specifications

Ros Roberts and Richard Gott

Evidence has a central role in all the new 2006 GCSEs in England. What do we need to teach and how might it be assessed?

Practical work at secondary level in England has a chequered history. At various times it has been illustrative in nature, then inductive (enquiry) in the Nuffield schemes, skills-focused and, of course and most recently, investigative. Its assessment has gone through similar phases. At one time a menu-driven practical was part of O-level syllabuses (for 14–16 year-olds). Then we had, at various times, investigations in Nuffield A-level (16–18 year olds), skills-based practicals and practical exams at O-level, none at all and, finally, the Science 1 (Sc1) coursework of the National Curriculum of recent years, which consisted of pupil reports of investigations. Whether we like it or not, this assessment drives the curriculum and its teaching to a greater or lesser extent. At present it leans towards 'greater' because of the high-stakes league tables with which we all live. As a consequence it is important that the changes to the key stage 4 (pupils age 14–16) science curriculum in England for 2006 are given very careful consideration.

In this article we explore the assessment model proposed by the awarding body AQA (the Assessment and Qualifications Alliance) for the new GCSE (this excludes the AQA 'applied' science

syllabuses, 4863 and 4861, and the ELC (Entry Level Certificate) syllabus), the examination completed at the end of compulsory schooling (age 16). The science education team at Durham has been working in close collaboration with AQA on the assessment of the 'How science works' section. The views here are, of course, ours alone.

First we discuss the criticisms that have been levelled at the current system, in order to generate an awareness of things to avoid in the revised arrangements. Then we examine elements of the new specifications and try to show how, by establishing a framework for considering evidence, we can consider what to teach and assess. Finally we look at the implications for the classroom and raise some issues that require constant attention if we are to avoid historical mistakes repeating themselves.

We start, then, with a look at the criticisms of the current system. They are many and varied but perhaps a quote from the House of Commons Select Committee (2002: 21) will set the tone:

The way in which coursework is assessed for GCSE science has little educational value and has turned practical work into a tedious and dull activity for both students and teachers.

The recent QCA (Qualifications and Curriculum Authority) report into coursework (2005a: 10) noted that science teachers:

referred to coursework as 'jumping through hoops' to maximise marks and regarded coursework as a poor educational tool.

It is very easy to criticise assessment methods. There are no simple solutions and rhetoric like the above, whilst making a newsworthy sound-bite, serves

ABSTRACT

This article considers the importance of evidence in the new key stage 4 (ages 14–16) National Curriculum for England section called 'How science works'. We specify the 'thinking behind the doing' as the basis for a procedural understanding and consider how this informs decisions about both what to teach and what to assess. We exemplify this in the context of one awarding body's response (that of the AQA) to the 2006 curriculum.

only to heat up the argument when what is needed is light. That requires us to try to identify the bit(s) of the current system that have given rise to these criticisms.

Shortcomings of previous coursework assessment

To begin with it is worth pointing out that Sc1 is a form of 'performance assessment'. That is to say, it is an assessment of how pupils actually carry out an investigative task. At least, that is its intention. We shall need to return to this later when we question whether this is an appropriate aim in the first place. For now, however, let us consider how 'performance assessment' can be carried out. Ideally the pupil, on his or her own, would carry out several open-ended tasks (for reasons of reliability and validity, discussed later) and their actions would be recorded somehow. Recording instruments might include such things as extensive checklists or video techniques. This is clearly a non-starter for GCSE assessment for any number of reasons. Within the pragmatics of the classroom, assessment has not been by direct observation of pupils' performance of investigations; so extensive use has been made of pupils' write-ups, which are usually submitted after opportunities to draft and re-draft the work. There are some doubts as to the extent to which these can be said to reflect pupils' performance (Baxter *et al.*, 1992), as well as worries about the extent of parental input and Internet plagiarism (QCA, 2005a), but these are not the issue for this article.

Assessment must be *valid* and *reliable*. Reliability and validity appear in this article in two contexts: the valid and reliable assessment of pupils and the valid and reliable evidence produced in practical work. Validity requires evidence that clearly bears on the problem: in the first of these contexts, if we are talking about assessment of understanding, say, the questions should not merely test recall. Reliable evidence is evidence that can be trusted – it should not rely on only one question in an exam for instance. For awarding bodies, understandably, reliability is a big issue. Pupils' futures depend on it. As a general rule, the more questions a pupil has to answer, the more reliable the final score is as an estimate of ability.

Gott and Duggan have argued elsewhere (Gott and Duggan, 2002) that awarding bodies have had two options for maintaining the reliability and validity of their assessment:

- Allow for lots of different practical investigations to be used for assessment. Some pupils do better

in some investigations than others (on account of the different contexts of the work, etc.) so they need to do lots of these tasks (maybe as many as 10) to 'iron out' these contextual effects (Solano-Flores *et al.*, 1999).

- Encourage only a small number of tasks, which are very tightly constrained. This doesn't reduce the context effect in principle, but a small number can be taught and practised, reducing variation and increasing reliability in practice.

Clearly the first solution is more *valid* than the second. There are many different ways of collecting and analysing evidence in science and by conducting several tasks there is the opportunity to reflect this diversity.

However, the first solution is another non-starter for performance assessment: it takes too long to do full reports of a large number of tasks – there would be no time left for teaching. So, in practice, the second solution was adopted in order to obtain reliable marks for the coursework. Awarding bodies achieved this in two ways:

- They emphasised the link to substantive (biology, chemistry and physics) subject knowledge (i.e. by giving marks to making predictions and links to 'theory' which credited ideas that are *already assessed* in the written papers).
- The tasks for assessment became 'routinised' – if pupils all 'do' more or less the same thing ('cloning' in Education Secretary Ruth Kelly's parlance (BBC News, 2005)), then it is easier to mark it reliably. These routinised tasks tended all to reflect just one way of working in science as defined by those 'standard' Sc1 investigations (Roberts and Gott, 2003).

So, Sc1 coursework has become acceptably reliable but these constrained tasks do little for its validity, or for its credibility amongst pupils and teachers.

A way forward – practical work as an 'end' or a 'means to (another) end'?

As noted earlier, the current system is based on a performance model. It makes the assumption that we hope pupils will learn to *do* science. This would enable them to work in a lab or in the field, use instruments and so on. Practical work, in this view, is an end in itself. But as we have seen, attempts to assess practical work in the 'performance' of whole investigations have serious shortcomings. Added to this, there is an increasingly widely held view that we should be educating pupils to read about and

challenge, as well as understand, science and how it affects their own lives. From this standpoint, it is far from obvious that teaching practical work as an end in itself is the best course.

However, if practical investigations are viewed as a way of solving problems in science, then we have a different way of looking at practical work. Practical work is not seen as an end in itself, but as a way in which ideas and skills are used to solve a problem. We already have reliable ways of assessing substantive ideas, so this leaves us to address the assessment of the procedural ideas (such as 'fair test', accuracy, repeat measurement, data interpretation, etc.) and skills. (By way of an aside, we have not considered practical work for teaching about substantive ideas such as forces, or rates of reaction, or feeding relationships. This is not because we dismiss such teaching, but rather that the assessment of these 'ends' is best carried out in the written examinations part of assessment, not in the coursework element.)

The substantive and procedural ideas and the skills that are required to solve practical problems (Roberts and Gott, 2004) are represented in the simplified model in Figure 1.

The AQA version of a solution to the assessment conundrum consists of two parts. It accepts that assessment should be of procedural ideas (of which more in the next section). But it also feels the need to include some assessment of skills, a sensible position to adopt as short skills assessments are easier to

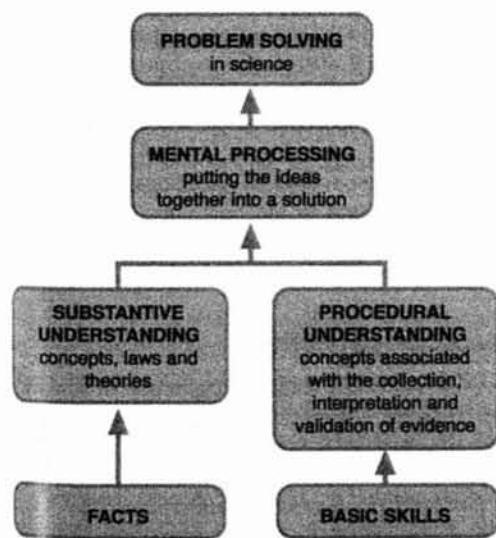


Figure 1 A simplified problem-solving model for science (based on Gott and Mashiter, 1991).

manage than assessment of whole investigations.

AQA is also constrained by the QCA requirements about teacher assessment and a wish to encourage a range of practical work in schools. This will be the focus of a later section. How it will work in the long term remains to be seen.

We are left, then, with two elements of teacher assessment:

- assessing procedural understanding;
- assessing skills (a performance assessment).

Before we can look in detail at the AQA solution and its consequences, we need to devote space to defining what we mean by 'procedural understanding'.

What are these procedural ideas?

The constituent ideas that underpin an understanding of scientific evidence have seldom been explicitly presented in texts and teaching resources (Roberts and Gott, 2000), although some more recent resources now target them (see Box 2, page 36). Some years ago, we published a tentative list of these constituent ideas, 'concepts of evidence', which, we suggest, go some way to defining 'procedural knowledge' and which underpin an understanding of scientific evidence (see Gott *et al.*, 2004).

The concepts of evidence have been validated against the ideas used every day by working scientists and technicians from industry and academic science research (Gott, Duggan and Johnson, 1999). Our research in Durham has attempted to delve below 'the things that scientists do' in a search for the understandings that are necessary pre-conditions of their work. Our aim was to determine what *underpinnings* must be *taught* so that pupils can be better prepared for the requirements of the workplace and have a critical awareness of science issues in everyday life.

The concepts of evidence attempt to specify the procedural knowledge base underpinning the collection and evaluation of evidence – the things we need to know to be able to judge validity and reliability, either from our own investigation or that reported by others. We have referred to this as 'the thinking behind the doing'.

If you imagine actually doing an investigation to solve a problem for which you don't know the answer, you need to make many decisions:

- Exactly what is the question you are going to investigate and how do you design a valid investigation to answer it?
- What do you need to measure to give you data, the reliability of which must be open to judgement?

- How will you take the measurements, and how many will be needed?
- What is the most valid way to analyse the data?
- What conclusions can be drawn, taking into account the reliability and validity of the investigation as a whole?

The concepts of evidence are *the ideas* used to make such decisions in an investigation. Figure 2 summarises these as a nested set of ideas. Essentially we need to be sure of the reliability and validity in each 'layer'. For example:

- For each datum we need to consider the quality of any reading taken.
- For a data set we need to consider whether sufficient repeated readings have been taken to capture the variation and enable us to trust the data.
- When seeking relationships between variables the validity of the design must be considered as well as the interpretation of the data.
- Comparison with other sources requires judgement on the validity and reliability of others' work.
- The reliability and validity of work must also take account of wider societal issues.

These ideas are integral to the planning and carrying out of practical investigations *with understanding* (rather than as a routinised procedure). Once we recognise that there is a set of ideas underpinning 'doing', then decisions about how to teach and assess these ideas can be made on a different basis. The discussion moves away from teaching and assessing practical work and moves on to ways of teaching and

assessing these procedural ideas, with practical work as just one of many ways in which they could be taught and assessed. We will address ways of teaching the concepts of evidence later and will also consider how AQA has addressed the assessment. But before this we need to consider the role of evidence in the new 2006 National Curriculum.

Why does understanding evidence matter?

Few would doubt the importance of evidence in science. Gott and Duggan have summarised these views in the *ASE guide to secondary science education* (Gott and Duggan, 2006: 189):

Science relies absolutely on evidence. This is its defining characteristic – theory must accord with reality. An understanding of scientific evidence matters for understanding science as a discipline but also, and arguably more importantly, because understanding evidence is essential for engaging with scientific issues in everyday life and for employment in science and science-related occupations.

One of the aims of the new 2006 key stage 4 science National Curriculum is the development of a critical approach to scientific evidence, another being the opportunity for pupils to acquire and apply skills, knowledge and understanding of how science works and its essential role in society. The current controversy about 'intelligent design' makes it even more important that the defining feature of science as a discipline lies in the absolute necessity to test ideas against reality by observation and measurement. If this cannot be done, it isn't science! That does not mean it is not important, but it is somebody else's issue, not ours.

The importance of pupils understanding evidence seems to be central to the new curriculum. Let us now consider how evidence fits into the new 2006 National Curriculum.

How science works

The QCA's new key stage 4 programme of study for 2006 includes a significant section entitled '*How science works*' (QCA, 2004), which must be included in all new GCSE syllabuses for science.

'*How science works*' is underpinned by ideas about evidence and, therefore, should enable pupils to gain an understanding about evidence that will be of use to them in their everyday lives as well as in science-based employment. The fundamental

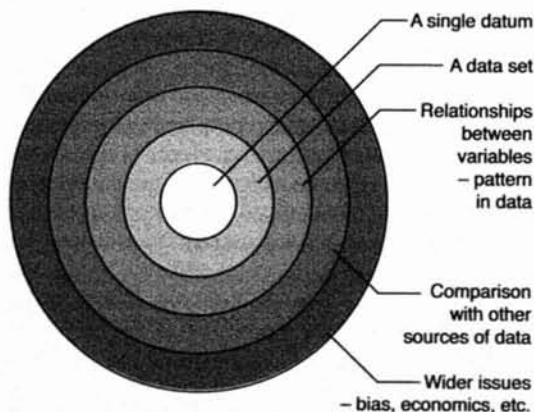


Figure 2 A summary of the concepts of evidence.

ideas of validity and reliability run throughout the specification, underpinning how scientific theories change over time, as well as the way scientists work.

'How science works' includes sections on:

- **data, evidence, theories and explanations**, which includes the collection, analysis, interpretation of data and the testing of ideas;
- **practical and enquiry skills**, which includes planning and data collection and the evaluation of data considering their validity and reliability;
- **communications skills**, which includes questioning, analysing and interpreting scientific information, and developing an argument and drawing conclusions;
- **applications and implications of science**, which includes how decisions in science are made and the role of the scientific community in validating scientific ideas.

So, we are in a position now to see the new 2006 curriculum as having a major component, 'How science works', underpinned by ideas about evidence. How does an understanding about evidence affect decisions about assessment in the new curriculum?

The assessment objectives (AO) for the new 2006 curriculum are shown in Box 1. These objectives are, perhaps necessarily, at a high level of generality. So we need to pin down what they might mean.

Let us now return to the two elements of the AQA scheme that make up the right-hand side of Figure 1 – procedural understanding and skills. The skills element appears just once, explicitly at least, in AO3a. Procedural understanding can be read into all the rest, depending on the interpretation of the words used. But it clearly plays an important and fairly obvious role in AO1b, AO2b and AO3b, c, and d. So we shall concentrate on those. How has the AQA scheme dealt with them?

AQA's approach to assessing 'How science works'

AQA has worked with us in the development of its science specifications and has used the concepts of evidence as the basis for specifying what is to be taught and assessed for 'How science works' (Hussain, Gott and Roberts, 2005). AQA has made these ideas explicit in its specifications (excluding the 'applied' science and ELC syllabuses) and has indicated how they will be incorporated in all its assessments.

In the AQA specifications, AO1 and the majority of AO2 are assessed by written papers or objective

BOX 1 The assessment objectives (AO) common to all awarding bodies (QCA, 2005b)

Assessment objective AO1: Knowledge and understanding of science and how science works

Candidates should be able to:

- a demonstrate knowledge and understanding of the scientific facts, concepts, techniques and terminology in the specification;
- b show understanding of how scientific evidence is collected and its relationship with scientific explanations and theories;
- c show understanding of how scientific knowledge and scientific ideas change over time and how these changes are validated.

Assessment objective AO2: Application of skills, knowledge and understanding

Candidates should be able to:

- a apply concepts, develop arguments or draw conclusions related to familiar and unfamiliar situations;
- b plan to carry out a practical task, test a scientific idea, answer a scientific question, or solve a scientific problem;
- c apply knowledge and understanding of how decisions about science and technology are made to different situations, including contemporary situations and those raising ethical issues;
- d evaluate the impact of scientific developments or processes on individuals, communities or the environment.

Assessment objective AO3: Practical, enquiry and data-handling skills

Candidates should be able to:

- a carry out practical tasks safely and skilfully;
- b evaluate the methods they use when collecting first-hand and secondary data;
- c analyse and interpret qualitative and quantitative data from different sources;
- d consider the validity and reliability of data in presenting and justifying conclusions.

tests. These will target questions at the substantive ideas in the syllabus as well as these procedural ideas. As AQA states, for example in its Science A specification (2005: 12):

Parts of questions may address procedural content, substantive content or blends of both ... candidates will be expected to apply procedural knowledge, understanding and skills in a wide range of contexts.

The centre-assessed unit, which forms 25 per cent of the overall mark, assesses mainly AO3 and some of AO2. Since changes to the teacher-assessed components of the AQA specifications are potentially the biggest change from current practice, we will focus on these here.

Centre-assessed unit

AQA considered that, for example in its Science A specification (AQA, 2005: 14):

The previous model of practical assessments based on 'investigations' has become a strait-jacket to practical activity in the classroom.

It has therefore changed its teacher-assessed component from performance assessment of practical work to the assessment of the underlying procedural understanding, *in specified contexts where pupils can do practical work*. In addition, there will be a continuous practical skills assessment.

The centre-assessed unit consists of:

- **An Investigative Skills Assignment or ISA.** The ISA involves the pupils carrying out practical work, drawn from a list of topics related to the curriculum and including fieldwork investigations, which they do in normal lessons. Following the practical work there will be a short externally set, internally assessed written test taken under controlled conditions which will include:
 - a written test of the pupil's understanding of the data and its collection from his/her own investigation (between 14–20 marks);
 - a written test of the pupil's understanding of other investigations relating to the same topic as the pupil's own investigation. This will include questions on the analysis and evaluation of data (between 14–20 marks).

These written tests, worth a total of 34 marks, will explicitly assess the procedural understanding specified by the concepts of evidence in the syllabus. Several tests will be available in

different contexts from the curriculum and the best mark obtained may be submitted.

- **A Practical Skills Assessment (PSA)** of the pupil's practical abilities, on a 'can do' basis, over the whole course, which will be done by the teacher (6 marks). The PSA covers normal class practical work in addition to practical work undertaken for the ISA.

We can see, then, that the model adopted by AQA is now largely one of practical work (as with other decisions about how to teach) as a means to an end. It is the 'end' of procedural understanding that is the main component of the assessment, with some credit being given to practical skills.

We have now described a series of ideas that could form the basis of a workable system:

- 1 **'How science works' needs careful definition.** The assessment objectives from QCA, and therefore incorporated within AQA and other awarding bodies' criteria, are at a high level of generality – too high to be anything more than a general guide.
- 2 **A Practical Skills Assessment (PSA).** The practicalities of assessment now revolve around assessment of practical skills (a hierarchical scale of implementation of practical work), which is not likely to be problematic to assess.
- 3 **An Investigative Skills Assignment (ISA).** A short, internally marked, written test, which is set in the context (expandable we hope) of practical work.

We can see that the concepts of evidence act as (part of) the knowledge base and become, as it were, a 'syllabus'. The awarding body now has the task of writing questions, using the concepts of evidence list as this syllabus. Having seen that specification of the knowledge base for evidence can affect decisions about how procedural understanding can be assessed, let us now consider how it could also affect how to teach.

What about teaching about evidence?

How often have we heard or made exasperated comments about pupils failing to understand what they have to do to evaluate their own investigations properly? How often have we been frustrated when pupils are asked to look at reports of others' evidence and all they tend to focus on are wider issues to do with the bias of the reporter or the funding body, whilst failing to get to grips with any of the issues to do with the actual quality of the evidence itself?

Perhaps it is because the pupils aren't really sure of the ideas necessary for making these judgements: they don't have a good grasp of 'the thinking behind the doing'. This is where the concepts of evidence are really useful: teaching them helps pupils to think critically about evidence. Our experience is that, once teachers have 'got' the idea of a procedural knowledge base, the problem becomes suddenly, and qualitatively, easier. If you know you have to get across the idea that the range and interval of a set of readings needs careful thought if the

underlying pattern is to be distinguished, then your own experience will come up with ways of teaching this based on a demonstration, a straightforward bit of teaching or a complete investigation chosen deliberately to target this idea.

In our experience with both teachers and examiners, having a clearly articulated set of ideas that can be taught and assessed has made it much easier to plan activities that address these ideas.

One of the challenges for teachers will be teaching these procedural ideas alongside the

Table 1 Exemplification of the sections from Figure 2 in two different subject areas relevant to the new GCSE science.

<i>Concepts of evidence associated with:</i>	<i>Examples of the sort of questions that can be answered by understanding the concepts of evidence. From the topic of:</i>	
	<i>Mobile phone masts</i>	<i>Blood pressure</i>
A single datum	A report provides a reading taken near the mast. Is this a valid measurement? Can the reading be trusted?	At the clinic the doctor has just told you that your blood pressure is 123/87. She was using a digital-readout sphygmomanometer. Is this a valid and reliable datum?
A data set	Have enough repeated readings been taken? How much confidence can be placed in the average value?	Joe Blogg's 'average' blood pressure is reported as 120/90 mmHg. What do you need to know about this 'average' to be able to judge whether Joe's BP is high?
Relationships between variables - patterns in data	In an investigation into mast energy levels and the incidence of headaches, have the right variables been considered and does the data show a relationship?	What would you need to take into account when designing an investigation into the relationship between body weight and blood pressure? Do the data collected show anything about the relationship?
Comparison with other sources	How do these results relate to other investigations to do with mobile phones and the siting of masts?	Are the data collected from 20 male patients at an obesity clinic typical of what is known from similar patients?
Wider issues	What issues concerning the people involved in the research and its reporting should be considered when evaluating the evidence?	Does it matter who collected these data? What factors should be considered?

biology, chemistry and physics specified in the syllabuses. The QCA has indicated that 'How science works' should be set entirely in the context of the substantive (biology, chemistry and physics) ideas of the syllabus (QCA, 2004). Teachers will need to be aware of the focus of any teaching activity and plan accordingly; this will not prove easy as many of the substantive areas identified for the new syllabus are ones not easily investigated at first hand.

The concepts of evidence underpin the decisions pupils have to make when conducting their own investigations in the context of the syllabus. But the same ideas underpin decisions being made about other people's evidence. Table 1 exemplifies the sections from Figure 2 in two different subject areas relevant to the new GCSE science: mobile phone masts and blood pressure. Both examples outline how the concepts of evidence can be used when evaluating others' evidence.

If we understand the concepts of evidence, we have got a basis for answering these sorts of questions. This knowledge base is therefore central to teaching pupils about evidence: there are ideas that can be taught. How the ideas are taught, whether through the use of practical work or using other teaching techniques, then becomes a decision for teachers. Some resources that suggest teaching activities are outlined in Box 2. Different types of teaching activities can be used to target different concepts of evidence. Some are illustrated in Figure 3 and Table 2 by way of example.

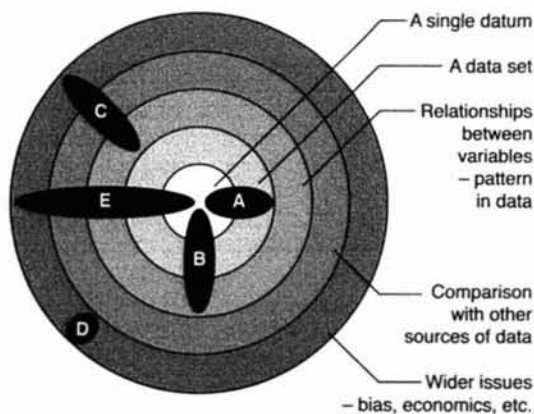


Figure 3 Different types of teaching activities that can be used to target different concepts of evidence.

BOX 2 Teaching resources for procedural understanding

Teaching the concepts of evidence through practical work and non-practical work

Although investigations (if carried out with understanding, rather than as ritual application of a routine) provide a way of applying the concepts of evidence, other types of practical work can also be used to explicitly teach some of the ideas, as exemplified in an earlier issue of *School Science Review* (Roberts, 2004).

Examples of different ways of teaching the ideas, including both practical and non-practical approaches, can be found in resources such as:

- Collins' *Science investigations* packs (Gott, Foulds *et al.*, 1997, 1998, 1999);
- ASE's AKSIS materials (Goldsworthy, Watson and Wood-Robinson, 1999, 2000);
- ASE's *Teaching secondary scientific enquiry* (Sang and Wood-Robinson, 2002);
- Folens' *Building success in Sc1 science* (Gott and Duggan, 2003);
- Key Stage 3 Strategy materials (DfES, 2005);
- The CASE materials (Adey and Shayer, 1994);
- *Thinking science* (Cheyney *et al.*, 2004).

These include non-practical activities, such as using text and media reports, different types of discussion and group-work activities or using ICT applications.

Concluding remarks

We began this article with a consideration of the problems of the current system. Any solution needs in our view, to:

- be responsive and dynamic – to avoid routinisation;
- have a light touch – to allow teachers to escape from some of the suffocating bureaucracy;
- encourage practical work in general and open-ended investigations in particular;
- return control of teaching to the teacher;
- allow as many different sorts of practical work as possible;

Table 2 Examples of the different teaching possibilities from Figure 3.

<i>Teaching possibilities represented by:</i>	<i>Concepts of evidence associated with:</i>	<i>How?</i>	<i>Example from topics in the GCSE science syllabus</i>
A	A single datum, a data set.	Questions targeted at data produced by the pupil doing a short practical measurement task or using data provided for them.	A class measured lichens as indicators of pollution. Using the data, do you think they've taken enough readings?
B	A single datum, a data set, relationships between variables and patterns.	Whole or parts of an investigation carried out by the pupil.	In an investigation to see whether reaction times are affected by caffeine, pupils could be asked to comment on the quality of the measurements they took, the design of their investigation and how the size and the representativeness of their sample affected their conclusions.
C	Relationships between variables and patterns, comparison with other data sources and wider issues.	Evaluation of a report using secondary sources.	In a report by scientists into the efficiency of wind turbines, analyse what the scientists did and have found and how this links with others' work.
D	Wider issues.	Analysis of a newspaper report.	How should the public respond to a report on the Internet entitled 'Contraceptive pill is not to blame for sex changes in fish say scientists from leading company'?
E	A single datum, a data set, relationships between variables and patterns, comparison with other data sources and wider issues.	A case study of a socio-scientific issue.	Provided with various pieces of information about an issue, such as data from research into emissions from an industrial chimney, reports from scientists involved, site maps to show where readings were taken, and newspaper reports about different pressure groups' reactions, pupils could be asked targeted questions from any of the concepts of evidence.
		A project	Pupils carry out their own investigation into nitrate levels in a stream and relate their findings, in an extended project, to others' work on eutrophication and local reports on environmental issues.

- contribute to the move to a more scientifically literate population.

Will the AQA arrangements do this?

On the positive side:

- There is a move towards practical work as a means of teaching procedural ideas – this emphasis is better suited to a science curriculum geared towards scientific literacy.
- In principle, the assessment is less constraining. If a good bank of written questions is built up, then it will not be so easy to teach to the test and more likely that teaching will concentrate on the underlying ideas.
- The list of contexts for practicals is still specified, unfortunately, but at least it is longer and has a broader range of tasks, including fieldwork, etc. It should be expanded so that the constraining, routinising effect is reduced.
- Again, in principle, the ISA and PSA are more efficient and therefore have a smaller backwash effect on the rest of teaching.

On the downside:

- Since only one practical is needed for assessment purposes it could result in even less practical work. However, before teachers had to do ScI assessment, they *chose* to do practical work – and more often and more interestingly than many feel constrained to do now. So, there seems to be no reason in principle why that should happen.

Ofsted could have a role in this, encouraging a range of stimulating practical work to teach the procedural ideas and skills.

- It could become routinised – so there will be a need to change the list of practical contexts and develop a bank of questions in just the same way as for substantive ideas. This will be important – the bank of questions must have new ones added (and old ones dropped) in an incremental fashion.

The existing practice of ScI became stultifying. The new curriculum, and in particular the emphasis on 'How science works', is a chance to do something different. For it to be successful, it is important to keep assessment in perspective: it needs to be efficient, valid and reliable and must have as little negative backwash as possible.

There is a major underlying shift in AQA's new curriculum: the assessment is of a set of ideas. Teachers will need to come to an understanding of:

- the ideas themselves (they are not as easy as some think!);
- how they underlie 'How science works', practical science (and engineering) and scientific literacy;
- how they are a critical part of the science curriculum.

It will take time but it will be an interesting challenge.

References

- Adey, P. and Shayer, M. (1994) *Really raising standards*. London: Routledge.
- AQA (2005) *GCSE Science A 2007/8 specification*. Manchester: Assessment and Qualifications Alliance.
- Baxter, G. P., Shavelson, R. J., Goldman, S. R. and Pine, J. (1992) Evaluation of procedure-based scoring for hands-on science assessment. *Journal of Educational Measurement*, **29**, 1–17.
- BBC News (November 2005) UK: Visited: December 2005. URL: <http://news.bbc.co.uk/1/hi/education/4457938.stm#>
- Cheyney, A., Flavell, H., Harrison, C., Hurst, G. and Yates, C. (2004) *Thinking through science*. London: John Murray.
- DfES (2005) *KS3 National Strategy: moving forward with teaching scientific enquiry and intervention*. CD-ROM. London: Department for Education and Skills.
- Goldworthy, A., Watson, R. and Wood-Robinson, V. (1999) AKSIS project. *Investigations: Getting to grips with graphs*. Hatfield: Association for Science Education.
- Goldworthy, A., Watson, R. and Wood-Robinson, V. (2000) AKSIS project. *Investigations: Developing understanding in scientific enquiry*. Hatfield: Association for Science Education.
- Gott, R., Duggan, S. and Johnson, P. (1999) What do practising and applied scientists do and what are the implications for science education? *Research in Science and Technological Education*, **17**, 97–107.
- Gott, R. and Duggan, S. (2002) Problems with assessment of performance in practical science: which way now? *Cambridge Journal of Education*, **32**, 183–201.
- Gott, R. and Duggan, S. (2003) *Building success in ScI science*. CD-ROM. Dunstable: Folens.
- Gott, R. and Duggan, S. (2006) Investigations, scientific literacy and evidence. In *ASE guide to secondary science education*, ed. Wood-Robinson, V. pp. 188–195. Hatfield: Association for Science Education.
- Gott, R., Foulds, K., Johnson, P., Jones, M. and Roberts, R. (1997, 1998, 1999) *Science investigations 1, 2 and 3*. London: Collins Educational.
- Gott, R. and Mashiter, J. (1991) Practical work in science – a task-based approach? In *Practical science*, ed. Woolnough, B. Buckingham: Open University Press.

- Gott, R., Duggan, S., Roberts, R. and Hussain, A. (2004) *Research into understanding scientific evidence*. Visited: March 2005. URL: <http://www.dur.ac.uk/richard.gott/Evidence/cofev.htm> (the website has downloadable versions of this fuller version plus a slimmed down option matched to the 2006 GCSE specifications by a group of teachers).
- House of Commons Science and Technology Committee (2002) *Science education from 14 to 19*. Third report of session 2001–2, I. London: The Stationery Office.
- Hussain, A., Gott, R. and Roberts, R. (2005) *AQA Specification for GCSE science 2006. Section 3.6: 'How science works'*. Durham: Durham University, School of Education.
- QCA (2004) *Programme of study: Science key stage 4*. Visited: December 2005. URL: www.qca.org.uk/downloads/10340_science_prog_of_study_from_2006_ks4.pdf
- QCA (2005a) *A review of GCE and GCSE coursework arrangements*. Visited: December 2005. URL: <http://www.qca.org.uk/downloads/coursework-report.pdf>
- QCA (2005b) *GCSE criteria for science*. Visited December 2005. URL: http://www.qca.org.uk/downloads/11881_gcse_science_criteria_apr05.pdf
- Roberts, R. (2004) Using different types of practical in a problem-solving model of science. *School Science Review*, **85**(312), 113–119.
- Roberts, R. and Gott, R. (2000) Procedural understanding in biology: how is it characterised in texts? *School Science Review*, **82**(298), 83–91.
- Roberts, R. and Gott, R. (2003) Assessment of biology investigations. *Journal of Biological Education*, **37**, 114–121.
- Roberts, R. and Gott, R. (2004) Assessment of Sci1: alternatives to coursework. *School Science Review*, **85**(313), 103–108.
- Sang, D. and Wood-Robinson, V. ed. (2002) *Teaching secondary scientific enquiry*. London: Association for Science Education/John Murray.
- Solano-Flores, G., Jovanovic, J., Shavelson, R. J. and Bachman, M. (1999) On the development and evaluation of a shell for generating science performance assessments. *International Journal of Science Education*, **21**, 293–315.

Ros Roberts lectures in science education and **Richard Gott** is professor of education at the School of Education, Durham University. Email: Rosalyn.Roberts@durham.ac.uk

British Physics Olympiad (BPhO' 06)
The 2006 Physics Olympiad Team
<http://www.pbho.org.uk>

The following students will form the 2006 British Olympiad Team to participate in the 2006 International Physics Olympiad to be held in Singapore, from 8th – 17th July.

Student	School
Peter Ford	Royal Grammar School, Worcester
Michael Mackay	Royal Grammar School, Newcastle upon Tyne
Jinyang Liu	Millfield School, Somerset
Matthew Norris	Wilson's School, Surrey
Jonathan Rees	St. Paul's School, London

The team was announced at a Presentation Ceremony, held at The Royal Society on Thursday 27th April. A full set of BPhO' 06 results are available on our web site under 'Winners'.

Information concerning the 2007 British Physics Olympiad Competition will be posted to UK schools in September 2006. The second round paper will be sent on Friday 3rd November 2006. Further information can be obtained from:

Dr. C. Isenberg, BPhO Secretary, Electronics Laboratory, University of Kent, Canterbury, Kent, CT2 7NT, Tel: 01227 823768, Fax: 01227 456084, email: c.isenberg@kent.ac.uk