SCORE
SCIENCE COMMUNITY REPRESENTING EDUCATION

2011 SCORE CONFERENCE REPORT

ASSESSMENT OF SCIENCE AT 14–19: IS IT FIT FOR PURPOSE?

Over 150 people interested in the assessment of school science attended the 2011 SCORE annual conference in London, on 11 February.

In its second year, the conference brought together academics, industrialists, teachers, educationalists and policy-makers to hear talks on the purpose of assessment and on related issues unique to science. The topic of standards was also on the agenda. The delegates had the opportunity to take part in discussions of what is a crucial and timely subject in light of the current review of the National Curriculum in England and the launch of Curriculum for Excellence in Scotland.

This report is an overview of the day.

Professor Graham Hutchings FRS chaired the conference.

ABOUT SCORE

SCORE is a partnership of organisations, which aims to improve science education in UK schools and colleges by supporting the development and implementation of effective education policy and projects. The partnership is currently chaired by Professor Graham Hutchings FRS and comprises the Association for Science Education, Institute of Physics, Royal Society, Royal Society of Chemistry, Science Council and Society of Biology.

SCORE's priority areas over the next four years are: the curriculum qualifications and assessment; the schools and colleges workforce; and the wider learning experience.

KEYNOTE SPEAKER

Professor Jonathan Osborne
The Shriram Family Professor in Science Education, Stanford University, California, US

Jonathan Osborne set the scene for the day, putting forward what he sees as four fundamental issues on assessment in science education.

- What counts matters, but not all that matters is counted. Persistence, curiosity, a critical disposition, a commitment to evidence as the basis of belief, and enthusiasm are central to being a good scientist. While there is no easy way of 'counting' such qualities, neither should any assessment system destroy them.
- The intentions of the curriculum are read from the assessment items, not the curriculum documents, so these must be of high quality.
- Assessments are the embodiment of what is valued, so it is vital to get them right.
- Critique matters as much as knowledge – knowing what is wrong is as important as knowing what is right.

According to Osborne, the assessment system needs to develop students' ability to be critical of information. The majority of assessment items are on recall, comprehension and application. What is also needed are ones that encourage analysis, synthesis and evaluation. An assessment system that is fit for purpose will have a spread of items that covers all six cognitive skills, and enables learners to demonstrate differences in ability, interest and aptitude in science.

KEYNOTE SPEAKER CONTINUED FROM P.1

One of the things we need to do is improve the flexibility of assessments.

Professor Jonathan Osborne

Osborne explained there are three kinds of knowledge needed to understand and work in science – content knowledge; procedural knowledge, ie concepts of evidence; and epistemic knowledge, ie ideas about science such as peer review and the nature of a theory in science. Most curricula are dominated by content knowledge with some measurement of students' procedural knowledge. There is too little focus on ensuring that students have some understanding of the role of models or hypotheses in science and the role of critique in improving ideas in science. This needs to be part of the assessment, he said. Osborne encouraged the delegates to think outside the box because future assessment mechanisms are likely to be more diverse as computer technology improves.

An international perspective on assessment is also important because the transnational tests – ie PISA[1] and TIMSS[2] – are a main concern of policy-makers, said Osborne. But it is important to understand what these tests assess because they are different from each other and from the UK curriculum. They place different emphasis on the different types of knowledge. The next round of PISA tests in science, in 2015, will be based on a new assessment model. One of the messages to bring to the review of the National Curriculum, said Osborne, is to think about assessment measures in framing what we want it to deliver. There is no point in having a wonderfully reliable system if what it is assessing doesn't count.

Issues raised by delegates

- The notion that what matters is just facts must be challenged because employers are looking for people who can think critically and apply their knowledge.
- In the trade-off between reliability and validity, what matters is validity.
- If assessments are to match the type of learning needed in science, the system will require more reliance on teacher assessment. But the public will need to be assured that this will be better than the existing system.
- The assessment system must inspire young people to do the subject. Part of the challenge of assessment will be to get pupils to think and engage with the subject.

... being asked to recall a piece of information is not particularly challenging nor stimulating, and doesn't engage more able students.

Professor Jonathan Osborne

[1] PISA - Programme for International Student Assessment
[2] TIMSS - Trends in International Mathematics and Science Study

THE PANEL DEBATE

The panel debate, chaired by Kate Bellingham, National STEM Careers Coordinator, gave delegates the opportunity to hear the views of end-users of the education system and whether the qualifications indicate what is useful to them.

Dr Nick Fletcher

Director of Education, School of Chemistry and Chemical Engineering, Queen's University Belfast

Fletcher had two main concerns. First, is grade inflation at A-level, which is putting up the entry requirements for university courses which, in turn, has a negative effect on widening participation. Secondly, the over-emphasis on 'teaching to the test' at the expense of understanding key but perceptionally hard areas of the curriculum is leading to a disconnect between what students are expected to know and an enthusiasm for their subject. Overall, he finds, incoming students have little in the way of skills-based knowledge and lack curiosity and engagement with their subject.

Dr Michael Porter

Admissions Tutor, School of Life Sciences, Bradford University

Porter uses the A-levels, and similar qualifications, to give him an indication of how good a student is, but he questioned whether this goes against what a student needs in their assessment to develop a strong interest in the subject. Porter finds students' focus on the exams is dampening their imagination and their ability to think critically about problems, especially when the latter are out of context. If there isn't a formula to punch numbers into, he said, they may struggle.

Julia Hatto

Scientific Technical Leader, Novartis Horsham Research Centre

Novartis looks for 'potential' in new recruits, said Hatto. Young people need the ability to use and develop the skills that they have. Too often, she said, they don't have the right approach to problem solving, they struggle to use their knowledge in unknown territory. They are too worried about being wrong, and poor at drawing conclusions from what they have observed. Hatto believes that teams in research industries need diversity to succeed. This diversity should be encouraged in the education system and reflected in routes to employment.

Cliff Billings

Head of Quality, ALcontrol Laboratories

Billings, who has been involved in the recruitment of chemists in various companies for the past 30 years, finds there has been too much of a focus on content and knowledge at the expense of the application of knowledge. It is the ability to apply knowledge, he said, that leads to competence. The assessment system, at school and university, he believes, should encourage problem-solving and creativity, both of which are crucial for the success of business and industry.

Key issues raised from the floor

- Assessment should capture students' abilities, knowledge and skills; written examinations alone are too narrow a way of doing this and other techniques, such as interviewing students, should be encouraged, where possible.
- Assessing a student's potential comes down to assessing higher-level skills such as analysis, evaluation and critique. It is disingenuous to have an exam system which doesn't give bright students the opportunity to show their understanding and skills.
- The publication of school league tables has led to an approach to teaching that doesn't allow students the opportunity to learn from failure, which is important in science.
- Students who learn with the sole aim of passing exams inevitably forget what they have learnt very quickly.
- Extended Projects at A-level can be used to develop high-level skills. Although they do attract UCAS points, they often do not form part of a university offer. When there is no pressure on the students to gain a particular grade, they are driven by their own motivation and interest in the subject.
- The science community needs to push the idea that teachers can be trained to be credible assessors.
- 'Grade inflation' is not restricted to the school system, universities have also been criticised for 'degree inflation' – a stark reminder that assessment needs to be looked at by all end-users working together.
- Any changes to the current assessment system should be based on research evidence rather than anecdote.
- Assessment should encourage students to see their careers as a path of life-long learning, during which they pick up extra layers of skills and knowledge with time.
- The most important end-users of education are the students, and the current system does not meets their needs. Focus on their needs, rather than criticising them for what they can't do, and they would be able to show their potential.

Examinations are a very narrow way of capturing the amount of information you need on individuals to make a good decision as to whether they are, or are not, qualified.

Dr Janet Brown

WHAT MAKES SCIENCE ASSESSMENT UNIQUE?

PART I: GRADING SEVERITY

Professor Robert Coe, *Director, Centre for Evaluation and Monitoring, School of Education, Durham University*

...whether one subject is harder than another very much depends on what grade you are talking about.

Professor Robert Coe

According to research[3] done by Professor Robert Coe, students who take physics, chemistry, biology and further maths, tend to achieve lower grades in those subjects than either they or other students achieve in such subjects as sociology, photography, film studies etc. But does this mean that grading is more severe in STEM subjects and so the same grade represents a higher level of difficulty? A positive answer might go some way to explain why some students aren't choosing to study STEM subjects at school and could have wider implications as to whether these subjects are taught in schools. Recent research by Coe found that an A* in chemistry and physics is worth about 20 UCAS points more than in law or sports studies. The science community needs to consider what, if anything, should be done. Coe put forward three options:

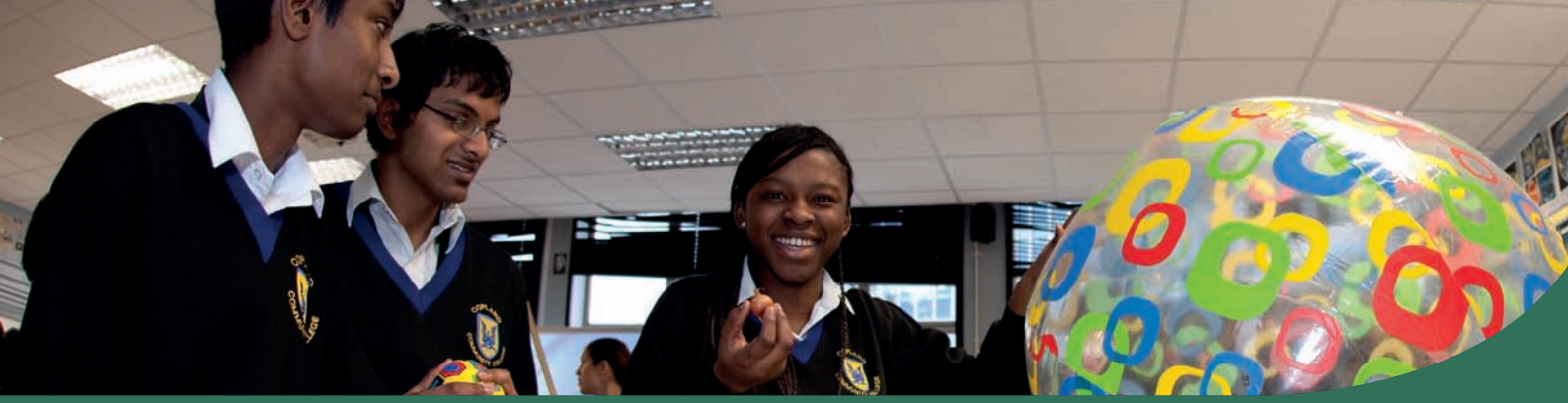
- do nothing – if comparability over time and across boards is more important, or if Higher Education Institutions (HEIs) require specific subjects for entry and don't treat subjects interchangeably;
- make grades comparable across subjects – but this would lead to STEM subjects being made easier and thus more difficult to discriminate at the top end than now;
- adjust the UCAS points tariff so grades are not comparable but UCAS points are.

According to Coe, the third option would solve more problems for STEM subjects than it would create and should be considered in light of the current review of the UCAS points system.

Issues raised by delegates

- The publication of school league tables gives schools an incentive to promote subjects in which it is easier to get an A grade, which could be addressed by adjusting the UCAS points tariff.
- The science education community needs to engage more with the press so that the right messages on the use and value of league tables get across to the public.
- If the perception of how difficult a subject is, and its UCAS points, change from one year to the next, public confidence in the system will be lost. However, this should not abdicate the science community from its responsibility to address a flaw in the system.
- There is evidence that the perception of STEM subjects as being harder than other subjects puts off students from less supportive and disadvantaged backgrounds.

[3] See **www.cemcentre.org/attachments/SCORE2008summary.pdf**



PART II: PRACTICAL WORK

Ros Roberts, *Senior Lecturer in Science/Biology Education, School of Education, Durham University*

We want lots of practical work, particularly open-ended investigations. They excite and engage pupils, and give them something to think about, which is what we should be encouraging.

Ros Roberts

Before deciding what aspects of practical work should be assessed, Ros Roberts reminded the delegates of why it is used to teach science. The list is huge – from learning how to observe, solve problems, and develop procedural understanding, to introducing scientific concepts, developing manual skills, following instructions etc.

The process of 'doing' practical work, ie observing, designing an investigation, making measurements and presenting the results, has in the past, explained Roberts, been assessed using performance assessment, which has led to some concerns about the reliability and validity of the assessment. Research finds that around 10 investigations are needed to assess reliably a student's ability to do an investigation. This would leave no time to teach, and so various attempts have been made to reduce the number of investigations needed. As a consequence, the range of investigations taught and assessed is narrow, and increasingly marks are given to 'theory', which brings into question the validity of the assessment.

As an alternative approach, Roberts and her colleagues at Durham looked at the ideas that underpin evidence. They concluded that open-ended investigations that are as wide-ranging and therefore as valid as possible should be encouraged in school science teaching. The assessment of procedural knowledge, whether formative or summative, said Roberts, should be done without disrupting the validity and scope of the science being taught. And there should be greater credit given for observation – 'seeing' from a science perspective. Difficulties associated with assessing open-ended investigations might be reduced by using video and computer technology. Or, elements of procedural knowledge could be assessed, such as selecting appropriate measuring instruments and allowing students to take sufficient readings to be confident in the data; written tests could then assess the underpinning ideas. Whether the assessment outcomes are the same for all learners, or whether they should be different depending on the progression route, needs to be debated.

Issues raised by delegates

- The Extended Project and teacher assessment could be used to address the failure of the current assessment system to recognise the importance of team work.
- There is a need to foster a professional learning community, through CPD provision, in which teachers from different schools come together to share ideas on what constitutes quality learning and how to assess it. This would nurture trust among teachers, which is often lacking, and ensure that assessment drives the curriculum in the right way.

THE APPARENT RISE AND FALL OF STANDARDS IN SCIENCE ASSESSMENT

TIM OATES

Group Director of Assessment, Research and Development, Cambridge Assessment

According to Tim Oates, there are standards associated with content, demand, attainment, standards over time, standards between specifications, between subjects and between types of qualifications, as well as teaching standards and standards of education. If you are talking about falling or rising standards, he said, you must be clear about what it is you are talking about, and clear about standard setting – ie getting the standards in a particular assessment right in the first place. Standards also need to be considered in light of the many mechanisms and factors that influence them – modularisation, accountability, the National Curriculum, the changing cohorts of higher education, the economy, changes to the structure and knowledge in particular subjects and so on.

What trends are evident in the English school science? Do they reflect a system that has 'gone to the dogs', or have there been massive improvements over the past 20 years? Oates says the evidence points to neither as being true. The system performance has remained static during a time of huge investment in education, and serial changes to the structure and number of qualifications.

...the notion of egalitarianism we have in our system has meant that people are absolutely opposed to routes through secondary education, and I think that is wrong.

So, can we learn from countries like Singapore or Korea that do well in transnational tests? Oates reasons that an understanding of how key factors combine differently in different settings can prompt us to look at our system to see what is possible through manipulation of these factors. But, he stressed, there are no assumptions that a single policy or dimension of another system can be imported readily into this country, with guaranteed benefit. The work of analysts such as Stigler and Stevenson shows that different pedagogic models operate in other, high-performing nations. As well as having primary phases which cover 'fewer things in more depth', differentiation is approached differently. In Singapore, classes move on to the next topic only when all the children have grasped the concept which is being addressed. If one child is struggling, teachers reason 'it is because the material has not yet been presented to that pupil in the right way' rather than 'they are of lower ability'. This contrasts with dominant models, in England, of progression, individualisation and ability.

Would England benefit from having a single awarding organisation? It would come at the expense of innovation in assessment which is possible through having a multiplicity of boards. A single body would, Oates warned, increase the dangers of un-mediated control in a country which has a long history of inconsistent policy directions. And it misses the point. To cater for different needs, the system requires different specifications in subjects such as maths, which raises questions of equivalence and standards over time. A single awarding organisation will not obviate the need for these equating mechanisms. And if they are in place and working, it is irrelevant as to whether the different qualifications come from one awarding organisation, or many.

DR JANET BROWN

Chief Executive, Scottish Qualifications Authority (SQA)

Dr Janet Brown provided an insight into the Scottish education system, which is separate and distinct from the rest of the UK. There are no league tables, and there is a single awarding organisation. The Scottish school curriculum is led by Government but, like its qualifications, she said, is developed in partnership with key stakeholders in Scotland.

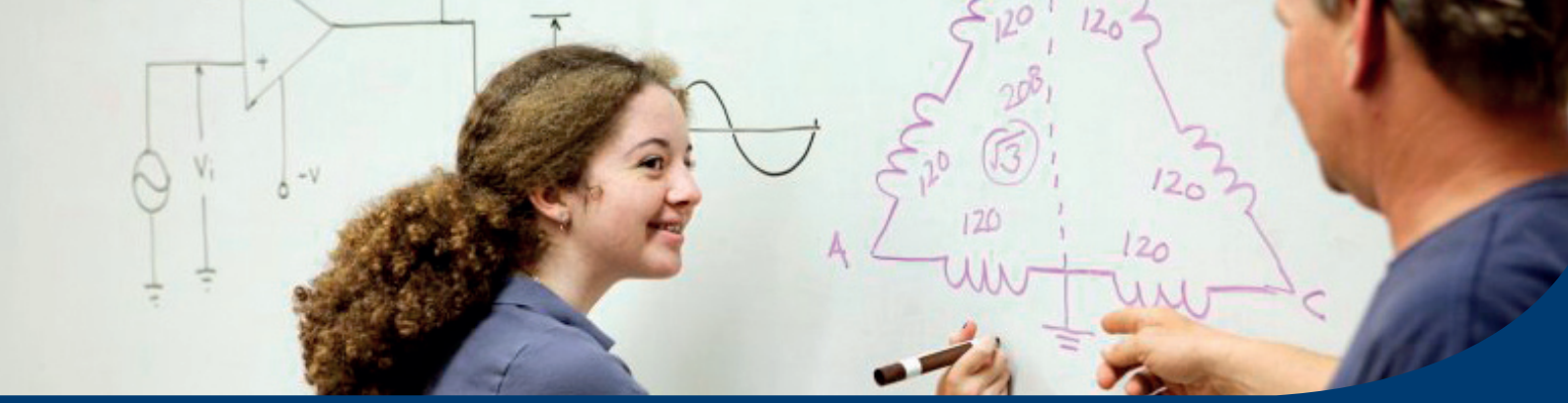
This year, Curriculum for Excellence[4] was launched – the culmination of six years' work to reform the school curriculum, from primary through to secondary education. The Scottish Qualifications Authority (SQA) is currently developing a new suite of qualifications that will be based on experiences and outcomes, and will aim to capture a broad range of knowledge, skills and abilities to support the various different progression routes.

Significantly, there will be more internal assessment to allow for variations in teaching and for the professional judgement of teachers. This represents a major change in culture for teachers because, alongside SQA, they will be responsible for maintaining standards for the qualifications. Teachers will be supported in this role by SQA. The authority will also be supporting quality assurance processes in schools and local authorities to create an environment where teachers can come together to share and understand standards. By allowing teachers to buy into the process, SQA believes will, in time, raise standards.

Issues raised by delegates

- A review of secondary science assessment should take a holistic approach and embrace primary science, as well as education beyond the age of 19.
- Assessment to support learners needs to be distinguished from assessment for accountability.
- Teacher assessment will work only if teachers are given appropriate CPD to develop their professional knowledge in this area.
- It is fine to teach to a test if assessment encourages broad coverage of the curriculum.
- A test should enable teachers to diagnose the nature of a student's learning.
- A single awarding organisation would eliminate the competitive, commercial market in which the current awarding organisations operate, and may reduce the problem of grade inflation.
- Innovation in the curriculum has come from external bodies such as Salters', Nuffield etc working in collaboration with schools and industry, and not the awarding organisations.
- The issue is not how many awarding organisations there are but the diversity and range of qualifications available to meet the needs of learners.

[4] Details of Curriculum for Excellence can be found at SQA's website (**www.sqa.uk/sqa/34714.html**).



SCORE'S WORK ON ASSESSMENT – IN BRIEF

DR ROSALIND MIST

SCORE Manager

Dr Rosalind Mist gave a brief excursion into SCORE's current activities in qualifications and assessment, stating its vision as 'an appropriate range of qualifications and assessment tools which should have a positive impact on the students learning'. In the past two years SCORE has commissioned research to support its vision in three areas.

- *GCSE science questions*. The research, led by Dr Colin Osborne, found there was a significant variation not only in the amount and level of mathematics required in GCSE examinations, but also in the amount and level of questions on How Science Works. There was also a lack of extended prose used.
- *The assessment of How Science Works in GCSE specifications.* The research, led by Andrew Hunt, found that there was an imbalance in the specifications of How Science Works between ideas covered by science in society and those that fell into the methods and nature of science. The research also identified exemplar materials to assess this part of the curriculum.
- *The nature and amount of mathematics used in assessing A-level sciences.* Led by the Institute of Physics and due to report in March 2012, the intention is to provide exemplars of A-level science assessments, and propose a set of mathematical skills required at each level.

The overarching aim of these projects, explained Mist, is to make SCORE more informed in these areas so that it can engage with policy-makers and the awarding organisations.

Currently, SCORE is looking at the role of Higher Education and learned societies in developing pre-university qualifications, as well as reviewing the way in which GCSE specifications are developed. Other areas of interest include the regulation and standards of qualifications, and the role of the awarding organisations.

It is crucial that the different parts of the science community work together to ensure that any changes to the curriculum contribute to improve learning rather than hinder it.

Professor Graham Hutchings

SCORE
SCIENCE COMMUNITY REPRESENTING EDUCATION

SCORE
6-9 Carlton House Terrace
London
SW1Y 5AG
tel: +44 (0) 20 7451 2245
email: score@score-education.org
web: www.score-education.org

The SCORE partnership, working together on science education: Association for Science Education; Institute of Physics; Royal Society; Royal Society of Chemistry; Science Council; and Society of Biology.