



GCSE Reform Consultation

SCORE's response to the Ofqual consultation

3 September 2013

Introduction

1. 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. The partnership is currently chaired by Professor Julia Buckingham and comprises the Association for Science Education, Institute of Physics, Royal Society, Royal Society of Chemistry and Society of Biology.
2. In summary:
 - SCORE supports the principle of reforming GCSEs, particularly if the opportunity is taken to ensure that appropriate qualifications are available in the sciences for the entire cohort, both those continuing to further study and those not doing so.
 - Many of the issues addressed by the consultation arise because of the competitive market between awarding organisations. SCORE is in favour of a system with a single awarding organisation, rather than multiple competing awarding organisations.
 - SCORE welcomes the proposal that the sciences will continue to be tiered. However, we are not convinced by the arguments in favour of an overlapping tiers model, and believe a foundation + extension assessment model would be more appropriate. Rigour and challenge should come primarily from assessment, not through the provision of different content for different groups of students
 - SCORE is pleased that practical work will continue to form a part of the assessment of students' final grades. However, further work is needed to establish the most appropriate way of managing this assessment to ensure both the validity and reliability of the assessment, and that students experience a range of practical activities in the classroom.
 - SCORE agrees with Ofqual that, given the extent of changes to GCSEs, a new grading system should be introduced. However, we would advise that, to avoid confusion, 1 should represent the highest grade, rather than 8.
 - It will also be important to see an end to the notion of a 'pass grade', the equivalent of the current grade C, since this distorts teaching and implies that those achieving less than this grade have failed and makes redundant any differentiation below D. SCORE recommends that an appropriate range of criterion-based assessment levels should be developed that cover all levels of ability, which should inform the number of grades available in the new system.
 - SCORE is pleased that there will be an annual review process for qualifications, and hopes that this will enable Ofqual to consult with subject groups, such as those within SCORE, on assessment and content.
 - SCORE is not convinced by the argument for reference tests, and fears they could stifle curriculum innovation.

General comments

3. A reform of this nature presents an ideal opportunity to re-examine the purpose and nature of qualifications at key stage 4. As SCORE stated in its response to the Department for Education consultation on key stage 4 reform¹, with all students being required to remain in some form of education or training until the age of 18, assessment at the age of 16 needs to have a clearly defined purpose, with the nature of the assessment being driven by that purpose. SCORE notes that the current consultation suggests that the reformed GCSEs should 'provide a basis for schools to be held accountable for the performance of all of their students'²; while this statement may represent the reality of how the results will be used, we are concerned that accountability will play such a major part in the design of the qualifications.
4. The sciences present particular challenges for reform. They are part of the core of the National Curriculum to key stage 4, but most students will not progress to further study in any of the sciences. Nevertheless, the qualifications need to ensure that those students who might progress to further study have an authentic experience of each of the sciences and are properly prepared for their next step. Ensuring effective provision for all students to allow for these different requirements presents a real challenge.
5. SCORE has produced guidelines for key stage 4 qualifications³ that outline recommended content for biology, chemistry and physics (plus earth science). These are structured around 'big ideas' for each of the sciences, and they will prove useful for those tasked with revising qualifications and curriculum for this educational stage. SCORE has also submitted detailed comments about the draft criteria for each of the science GCSEs in its response to the Department for Education consultation⁴.
6. Many of the issues addressed by the consultation arise from the fact that England has competition between awarding organisations. There is nothing within the market that maintains standards or encourages high quality assessments (and therefore high quality learning experiences in preparing for them). Instead, the market is driven by price, accessibility and outcomes. It is apparent that the market (combined with school league tables) has led, over the years, to a drop in standards and a fall in the quality of assessments.
7. It is also apparent from this consultation that the current system entails regulatory requirements to ensure inter-board comparability, which distort decisions about areas such as grading, so that decisions are made on the basis of those requirements rather than what might be best for students.

While SCORE did not favour the specific franchising process envisaged by the Department for Education, we would still favour a system that removed competition between awarding organisations from the provision of qualifications. That said, a key

¹ The response can be found at <http://score-education.org/media/11406/score%20ks4%20response.pdf>

² Ofqual consultation document, p.8

³ <http://score-education.org/media/12525/ks4%20guidelines%20final%20version.pdf>

⁴ The response can be found at <http://score-education.org/news/consultation-responses>

characteristic we would like to see of whichever system is in place is its ability to foster curriculum innovation.

Tiering

8. SCORE welcomes the proposal that the sciences will continue to be tiered at GCSE; we believe that tiering allows for more appropriate assessment for the full range of students in the science subjects, allowing stretch and challenge at the top of the ability range, and allowing those nearer the bottom of the range to demonstrate their knowledge effectively.
9. SCORE notes that Ofqual is in favour of an 'improved version' of the overlapping tiers model, but does not state how it would be improved. SCORE would welcome more clarification on what Ofqual propose. We do not feel that familiarity with the model is a good reason for making no change, and propose that the foundation + extension assessment model would represent an improvement on the existing model, for both the double award in science and in the separate sciences. The reasons for this are outlined below.
10. While there may be some areas of content from across biology, chemistry and physics that are more suited for high-ability students, rigour and challenge should come primarily from assessment, not through the provision of different content for different groups of students. A foundation + extension assessment model, where all students take a foundation paper while more able students can opt to take an additional paper to access higher grades, would allow the content to be broadly the same for both groups. Additional differentiation in the extension paper would be available through more challenging and higher order questions.
11. Many of the objections to tiering arise because it can force schools to make decisions about choice of tier too early. If content is broadly the same, decisions can be taken later. This would also facilitate progression, as all students would be more likely to have studied the essential content needed to take subjects further.
12. Where possible, tiering should be structured to avoid students being penalised for attempting a higher tier, missing out on a grade that their performance would allow them to have achieved if they had only taken a lower tier. Such tiering will lead to schools making decisions about entry to papers based on what they believe to be safest for the school, rather than best for the student. This effect is exacerbated if school accountability systems measure achievement of a particular grade (for example, a grade C, as in the current accountability measures), rather than valuing all achievement. The foundation + extension assessment model would help to overcome these issues, as it would mean that the decision on entry is delayed until schools have a clearer idea of students' abilities, and there would be less of a 'penalty' for inappropriate entry.
13. It is also important that there be no price disincentive for schools to enter students for higher or additional tiers, so awarding organisations should cost their qualifications with this in mind, ensuring that any additional costs associated with further assessment are not borne by schools.
14. Whichever model is decided upon, it is important to give proper thought to the content and assessment of the foundation tier. As we state above, science is a core subject in the National Curriculum so the content of the foundation tier needs to be appropriate for

the majority of students, and needs to reflect the National Curriculum programme of study for key stage 4 (though will also include additional content). It should not be a 'cut down' version of the content of the extension or higher tiers, but a coherent and properly structured qualification in its own right.

Assessment

15. SCORE is pleased that some element of practical assessment has been retained, since this recognises the centrality of practical work in the study of the sciences. Practical work in science serves a number of purposes⁵, including the provision of first-hand experiences to support the learning of science content. In addition, students should be expected to be proficient in a number of practical and manipulative techniques, and to develop scientific ways of reasoning. Distinction needs to be made between the various purposes of practical work in science, and subsequently how each of these aspects should be assessed.
16. The development of hands-on practical skills, and the ability to manipulate equipment, is part of what it means to be a good scientist and it would be preferable if they were assessed. It is hard to see how these techniques would be assessed indirectly, so there needs to be some form of direct assessment. Any such assessment must be valid and robust. If the assessment is designed to provide differentiation, then the differentiation must be reliable with a wide spread of marks. However, consideration should also be given to a model where direct assessment is not intended to provide differentiation, but rather to check whether or not a student is competent in a given practical skill.
17. We also expect students to develop good procedural understanding of how to plan, carry out and evaluate aspects of practical work and analyse results. These aspects can be assessed through the written examination in such a way that students are likely to perform better if they have acquired that procedural knowledge through doing practicals. It is up to the awarding organisations to set good questions to ensure that happens.
18. The reform of GCSEs offers an opportunity to think creatively about how practical skills and knowledge can be effectively assessed. Awarding organisations must be involved in that reform, and it is essential that assessment modes are piloted before being introduced nationally. SCORE is supportive of the proposal that teachers should be involved in assessment, but appropriate safeguards need to be put in place to ensure the validity and fairness of the assessment.
19. Awarding organisations also need to take into account in their planning the fact that proper assessment of practical work, and of the sciences more generally, has costs attached, and that examiners and moderators need to have appropriate expertise. Pressure to reduce costs, resulting in lower quality assessment, is a major shortcoming of a competitive market between awarding organisations.
20. SCORE has recently conducted research into the resourcing of practical work in schools⁶, which revealed worrying gaps in the provision of appropriate equipment and

⁵ See SCORE's 2008 report on practical work in science: <http://score-education.org/media/3668/report.pdf>

⁶ <http://score-education.org/policy/curriculum/practical-work-in-science>

access for carrying out practical science. In addition, over 68% of respondents reported that controlled assessments were a major driver of the amount and type of equipment resourced for practical science. However practical work is assessed, awarding organisations should ensure that they only specify equipment that schools should already have, or would be able to continue to make use of in the future. SCORE's benchmarks would be a useful guide to identify such equipment.

21. Whatever assessment scheme is proposed (for all aspects of practical work), there needs to be a common approach by all awarding organisations. They should not compete on their practical assessment scheme.
22. It is proposed that direct assessment of practical work will account for 10% of the overall assessment marks, with a further 10% coming from indirect assessment. However, practical work accounts for 25% of the marks in the current GCSEs, so the reduction could be taken by some as a downgrading of its importance. SCORE would like to see further protection for ensuring that students do carry out practical work in the classroom. More significant than the specific weighting is the importance of ensuring that all students continue to have a proper experience of practical work, and that its central place in learning science is recognised in the assessment and certification of qualifications in the sciences.
23. The sciences are not among the subjects that currently have marks explicitly allocated to spelling, punctuation and grammar (as they are in specifications in English literature, geography, history and religious studies). However, some specifications do include marks allocated to quality of written communication, and students cannot access higher marks in these sections with poor SPAG. It is important to ensure that a percentage of marks continues to be allocated to the quality of writing, as the ability to express scientific ideas clearly is an important part of scientific learning, and this should be recognised in assessment. This is particularly the case given that there is the intention to include more extended-answer questions, which SCORE supports.
24. SCORE welcomes the greater emphasis on synoptic assessment. We particularly welcome the opportunity for awarding organisations to include different types of assessment in their question papers, such as structured and unstructured questions and multiple choice, and would welcome greater regulation on the use of assessment that is appropriate for the subject.

Grading

25. SCORE agrees with Ofqual that, given the extent of the changes to GCSEs, it would be appropriate to signal those changes with a new grading system. Moving to a numeric system seems sensible. It is worth noting that the same argument will apply to A-levels when they are reformed.
26. However, we do not accept the reasons for having 8 as the highest grade, and propose that 1 should be the highest. This is for a number of reasons:
 - Where numbers are used elsewhere as a rank order, 1 is usually the highest, and to do otherwise could be confusing;

- Starting at 8 to leave 'head room' to add further grades if necessary admits the likelihood of grade inflation; it would be more appropriate to introduce measures that militated against such a need. Introducing higher grades would both de-value grades awarded in previous years and cause confusion about the level of performance a particular grade represents;
27. There is also a danger that, if eight grades continue to be used (8 to 1 or A* to G), people will look for, or assume that there are, equivalences between the new and old systems. It is therefore advisable for the new grading system to have a different number of grade points. Even if there are no explicit equivalences between the old and new grading systems, there will need to be clear communication about the levels of performance indicated by each of the new grades. This will be particularly acute for the first few cohorts going through the new system, who may also have qualifications using the previous grading system.
 28. Under the current grading system, grades E, F and G comprise less than 15% of the cohort, but 38% of the grading structure, which suggests a need to look again at the way grading is carried out, so that the grades awarded make use of the full range available. However, it will be important to retain an appropriate amount of discrimination at the lower end of the ability range, to ensure that all candidates have ample opportunity to demonstrate their abilities, and a grading system that recognises that achievement. SCORE recommends that an appropriate range of criterion-based assessment levels should be developed that cover all levels of ability, which should inform the number of grades available in the new system.
 29. SCORE would also like to see an end to the notion of a threshold 'pass' grade, the equivalent of the current grade C; the emphasis on achieving this threshold has distorted teaching and implies that grades achieved below this level are effectively fail grades. This shift in emphasis needs to be reflected in the accountability measures.
 30. Clarity is needed on how grades will be awarded for the double science GCSE; it might be useful, for example, to provide information on students' performance in the three separate science elements and working scientifically, as well as any synoptic elements within the assessment. Such information would also be useful to provide for students taking the sciences separately. SCORE understands the concerns around the impact that providing subsidiary information may have on assessment design, but such information is valuable to students and schools, so we feel it is worthwhile exploring ways in which it can be provided.
 31. SCORE is concerned about the suggestion that only 'broad comparability' can be achieved in the grades awarded by different awarding organisations. As we have stated above, a market-driven system in qualifications provision has a number of disadvantages, of which the lack of comparability between grades awarded by different awarding organisations is perhaps the most serious. It is of concern that this apparent lack of comparability is also the main driver in the decision to restrict the information provided to candidates on their performance, rather than what is most beneficial to students, schools and other stakeholders.

Regulation

32. SCORE is pleased that there will be an annual review process to evaluate GCSEs and provide a mechanism for expert feedback. We hope that this will enable Ofqual to consult with appropriate subject groups, such as those in SCORE, to ensure that qualifications are regulated effectively at subject level, as well as awarding organisation level.
33. SCORE is sceptical about the notion of 'market and regulatory incentives' being used to improve awarding organisation performance. It is not clear how effectively the market actually operates, since schools choose qualifications for a range of reasons, including maximising their performance in league tables and avoiding undue financial cost (for example needing new textbooks), that have little to do with the quality of the assessment. In addition, the users of qualifications – higher education, employers, candidates themselves – are not part of this market, although they are key stakeholders.
34. SCORE is concerned about the use of reference tests to assist with the maintenance of standards over time:
- It is unclear why they are needed, other than to mitigate the difficulties of setting standards in a competitive market, in which case it may be the system itself that requires attention, as we discuss in paragraph 6 above.
 - They could stifle curriculum innovation if awarding organisations need to maintain a relationship between their own qualifications and the content of the reference tests over a period of years or even decades. This could be a particular issue for the sciences, the content of which can change dramatically over time.
 - If taken at the same time as GCSEs, they would also place an additional burden on students at a time when they are preparing for their examinations.
 - We are also concerned that if the questions for the reference tests are kept the same each year, the average mark is likely to increase without any rise in standards.