



Reforming Key Stage 4 qualifications

SCORE's response to the Department for Education consultation

10 December 2012

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 Graham Hutchings FRS and comprises the Association for Science Education, Institute of Physics, Royal Society, Royal Society of Chemistry and Society of Biology.
2. In summary:
 - If new qualifications are to be put in place for the end of Key Stage 4, they must have a clearly defined purpose. This must be considered alongside any changes to the accountability structure for schools. SCORE is developing guidelines for science qualifications at key stage 4 and will make these publicly available in February 2013 to help inform specification development and selection.
 - SCORE is very concerned that the EBacc may only require a student to study two of the three science disciplines. This would be detrimental both to uptake in the sciences beyond 16, and to the education of all, since it would limit students' exposure to one or more of the sciences. Students should be required to study all three sciences, even if the EBacc can be awarded for a pass in only two of them. SCORE would like to see further consideration given to the provision of combined science qualifications at Key Stage 4, in addition to separate sciences.
 - SCORE does not believe that single-tiered assessment can accurately and fairly measure performance across the whole ability range, and therefore we would support the retention of some form of tiered assessment.
 - SCORE would like to see the quality of assessment placed at the centre of any reform. It is the content and quality of assessment that drives teaching in schools, rather than the specifications. It is imperative that assessment is well designed to ensure full curriculum coverage and progression in the development of skills. Many of the problems identified in the consultation document could be addressed by improving the quality of the assessment items at Key Stage 4, without making wholesale changes to structure or content.
 - Practical work is an essential part of science and must be safeguarded in any reform. Practical work must therefore be assessed as part of any Key Stage 4 qualification in science subjects; without assessment, there would be a severe risk that practical work will no longer be taught in science lessons.
 - SCORE cannot support the franchise model described in the consultation document. It limits the opportunities for awarding organisations to engage with subject experts during the development process and prevents the provision of alternative specifications and curriculum innovation. SCORE would, however, be supportive of the proposal to have subject suites of qualifications offered by single awarding organisations, which could be achieved either by a single national awarding organisation, or by awarding franchises to awarding organisations rather than individual specifications. Such a process should include explicit requirements for investment and research to improve the quality of assessments, and engagement with stakeholders and subject experts.

- SCORE believes the timeframe proposed for the introduction of the new qualifications is unworkable and risks destabilising the examinations system. Instead, more time should be taken to consider reform of the education system as a whole.
- The Department for Education should consider any reform of Key Stage 4 qualifications alongside the reform of the National Curriculum, A-levels and school accountability, to ensure the resulting system is coherent and allows for progression from one stage to the next.

Purpose of reform

3. Assessment has a range of different purposes, including accountability, certification, progression to the next stage of education and differentiation between candidates for selection. Assessment systems operate most effectively when they have a primary use for which the assessment has been designed, so there does at least need to be a prioritisation of purposes, in order to design a system that is fit for purpose¹.
4. SCORE is concerned that there is no clearly defined rationale in the consultation document for the retention of high-stakes, externally marked qualifications at 16, given that all students will shortly be required to remain in education or training until the age of 18. It is important to define the purpose of assessment at this stage, since this should be the primary driver of the proposed reform; it could still be used to provide a record of achievement for those choosing not to take particular subjects further, and provide diagnostic information for students and teachers about where their strengths lie. However, without a clear understanding of what is to be achieved by assessing students at this point in their educational lives, it will be difficult to judge the efficacy what is being proposed.
5. The consultation document attempts to divorce accountability from qualification development; we understand accountability will be discussed in a separate consultation. However, it is impossible to consider the two issues separately, since we know that using assessment for this purpose has a distorting effect on the way subjects are taught. We therefore recommend that the two issues are considered together. We recommend the consultation on accountability should explore whether the EBacc has value as an accountability measure at 16, and its suitability or otherwise for progression to a range of destinations post-16.
6. SCORE welcomes the opportunity to improve Key Stage 4 assessment, but if it is to be successful the Department must consider the whole education landscape. There must be consistency and communication between the current reforms of the National Curriculum, Key Stage 4 and A-level to ensure that the resulting system is coherent and allows progression from one stage to the next. The pace of change and the number of consultations also makes it difficult for stakeholders to maintain a sense of a holistic education system being designed.

¹ see Newton, P (2007), *Clarifying the purposes of educational assessment*, *Assessment in Education* 14:2, 149-70 for a discussion of the multiple purposes of assessment, and the importance of clarifying or prioritising those purposes.

7. SCORE is also concerned that the setting of criteria for assessment has a major impact on the quality of assessment, yet the process for setting assessment criteria has received insufficient attention. There must be clearer linkage between Key Stage 4 national curriculum development, the setting of subject assessment criteria and the writing of specifications.
8. A significant failing of the current Key Stage 4 assessment system is the quality of assessment items, and overall schemes of assessment, but these areas are not addressed in the consultation. High quality assessment is required to ensure the resulting qualification remains authentic to the subject, enables progression and drives high quality teaching and learning. Many of the problems identified in the consultation document could be addressed by improving the quality of the assessment items at Key Stage 4, without making wholesale changes to structure or content.
9. SCORE disagrees with the proposal to name the new qualifications English Baccalaureate Certificates. The term Baccalaureate is used internationally to refer to qualifications taken at 18, so using the term for a qualification at 16 would be inconsistent.

Development of Key Stage 4 qualifications

10. SCORE has on a number of occasions outlined the characteristics of an effective examination system². Using these characteristics as a benchmark, we support the notion of a single awarding organisation for the development and provision of the Key Stage 4 science qualifications. We see many benefits to this model, for example: money previously spent on marketing the qualifications could be invested in assessment development; the removal of competition should end the “race to the bottom”; greater transparency for users of the qualification; and it would offer the opportunity to provide a coherent suite of qualifications in the sciences that complement each other while each retaining their own rationale, something that is hard to achieve with multiple awarding organisations.
11. However, we do not support the current proposed franchise model, for the following reasons:
 - Franchises are to be awarded on the basis of a competition between submitted specifications. There are a number of problems with this:
 - i. There will be limited opportunities for awarding organisations to engage with subject experts during the development process, both because of the time available, commercial sensitivities and because all the development needs to be done upfront, in advance of the competition. Once specifications have been selected, there will be little or no opportunity to make changes. This will have an impact on the quality of the specifications.

² SCORE’s response to Education Select Committee inquiry on examinations: <http://www.score-education.org/media/9364/ed%20select%20exams.pdf>

- ii. It is not clear how it would still be possible for awarding organisations to offer alternative specifications such as those developed in partnership with the Nuffield Foundation. This would have very negative effects on curriculum innovation.
 - iii. It is not clear what criteria will be used to select the winning specification, but there is a danger that the franchises could be awarded for reasons that could have negative educational repercussions. It would also be important for the bids to be judged by those with the appropriate expertise, in both subject knowledge and assessment.
- A franchise model, with franchises to be re-tendered every five years, risks destabilising the examination system, since it could damage awarding organisations' capacity to run their operations, and schools will still be dealing with multiple administrative systems (unless the same awarding organisation is awarded all the franchises). A five-year contract is not long enough to allow awarding organisations to judge the effectiveness and impact of the specifications they are running, and would mean that they would need to start working on their re-submissions within a year or two of being awarded the contract, and before the first cohort had sat the assessments. It would also mean considerable disruption for schools, who would need to make changes to schemes of work and other material every time the specification changed. Should the franchise go ahead we would propose an eight-year contract instead, which would allow four full cohorts to take the qualifications, and allowing time for investment and reflection.
 - It is also not clear whether awarding organisations will have the capacity to compete for future franchises in subjects they do not win in the first tender, since they are likely to have to make staff cuts in subject areas where they are not the winning EBC provider. This could also have an impact on how awarding organisations choose to develop other qualifications such as A-levels. This again highlights the need to consider Key Stage 4 and A-level reforms together.
12. Instead of the proposed franchise model, SCORE would propose a single awarding organisation, as is the case in Scotland and Northern Ireland. This would remove the negative impact of competition from the qualifications system, while putting in place a structure that would allow curriculum and assessment innovation and expertise to be used to greatest effect. It would also facilitate greater coherence between the sciences and mathematics, and indeed other subjects, since they would be developed by the same awarding organisation. Alternatively, franchises could be awarded on the basis of a more rounded submission from awarding organisations, rather than individual specifications, a process that would solve many of the issues identified in point 10 above. Such a submission could require awarding organisations to explain how they would invest in curriculum and assessment innovation and development.
13. SCORE is pleased that subject criteria will remain to guide the development of Key Stage 4 qualifications. The criteria must offer more guidance than is currently the case, describing the level of depth required in addition to the core knowledge, and giving some guidance on suitable assessment requirements. SCORE would advise that subject experts write the criteria, in consultation with the relevant subject communities (in the

sciences this would be the constituent members of the SCORE partner organisations) which should build on the scientific ideas developed through the National Curriculum in earlier years. It would also be important for there to be an appropriate timeframe for this work. Currently, failings in subject criteria are having a damaging effect on the quality of science assessment at Key Stage 4.

14. SCORE would also propose that developing the criteria for Key Stage 4 qualifications could be part of the remit of the national subject committees that we have proposed in our response to the Ofqual consultation on A-level reform³. The committees would be made up of people with expertise from a variety of stakeholder groups, including universities, employers and teachers, and would have three main responsibilities: to define the criteria from which awarding organisations design their A-level qualifications; to decide whether qualifications meet those criteria, before the qualifications can be accredited by Ofqual; and to scrutinise a sample of assessment material on an ongoing basis (as opposed to just when the qualification is presented prior to accreditation) to ensure sufficient curriculum coverage and level of demand. Giving these committees a role in the development of Key Stage 4 qualifications would ensure that there was clear progression from one stage to the next.
15. It is not clear from this consultation how and when criteria for Key Stage 4 qualifications will be developed. It is also not clear on what basis Ofqual and the Department for Education will make their decisions on accrediting specifications and identifying the successful bidder. This lack of guidance is unhelpful and particularly worrying given the awarding organisations are already developing specifications and, to the best of our knowledge, intend to submit specifications to Ofqual for accreditation in May 2013.
16. SCORE is therefore committed to developing guidelines on the characteristics of appropriate Key Stage 4 science qualifications. These guidelines will identify the purpose of science education, the principles awarding organisations should adhere to in developing a science qualification and the content of Key Stage 4 science assessment. They will include:
 - overarching principles in biology, chemistry and physics
 - the assessment objectives for practical skills in science
 - the mathematical requirements all 16-year should be able to apply within a science context.

SCORE will make these guidelines publicly available in early February to advise awarding organisations in their specification development and to advise Ofqual and the Department on the criteria they intend to use to appoint a successful bidder.

17. SCORE is concerned about the fact that very different processes are proposed for the development of Key Stage 4 qualifications and A-levels. At A-level the proposed development process is decentralised, Government is not accountable for the qualifications, there are no subject criteria and the qualification will exist in a free market. At Key Stage 4 it is proposed that there would be strong centralised control (with the Secretary of State having ultimate responsibility in appointing the franchise holder for a

³ SCORE's response to the Ofqual consultation on A-level reform: <http://www.score-education.org/media/11200/alevel.pdf>

suite of qualifications), there would be subject criteria, there would be a route back to Government and the qualifications would only be offered by one awarding organisation. We can see no rationale for these differences and would strongly advise that there is a single model for qualification development to avoid confusion among the users (higher education and employers) and lack of confidence in the process.

18. Furthermore, however quality assurance of the new qualifications is carried out (at Key Stage 4 and at A-level), the roles and responsibilities of all organisations involved – Department for Education, Ofqual, awarding organisations and national subject committees – will need to be fully defined and transparent.
19. SCORE agrees that support provided by awarding organisations to schools and colleges must not encourage a narrowing of the curriculum. In particular, SCORE would advocate breaking the link between awarding organisations and publishers. This would free schools to use the support material and books that they felt best suited their students, and based on the quality of the materials concerned, without feeling constrained to use the material endorsed by the awarding organisation in a particular subject.
20. Awarding organisations should not be tasked with ensuring high quality teaching and learning. Their responsibility is for developing innovative, engaging and appropriate specifications and high quality assessment, which, if designed appropriately, should encourage teaching and learning that ensures a real understanding of the subject and demonstration of that knowledge.

Science at Key Stage 4

21. SCORE supports the teaching of three separate sciences with their own identities from the start of Key Stage 4, regardless of the assessment put in place at the end of the Key Stage, in order that students are able to make informed choices about progression. However, science education does much more than ensure a future generation of STEM graduates; it serves to ensure all young people are scientifically literate. Given that a large proportion of the population will end formal science education at 16, the Key Stage 4 science assessments must consider the needs of these students in addition to the needs of those who will progress in the separate sciences. We therefore do not feel it would be appropriate to offer only separate science qualifications; there must be the option of a combined science assessment that allow students to progress in the sciences should they wish to do so, and include elements of all three sciences. This option would provide greater flexibility in the Key Stage 4 school curriculum.

We recognise that more thought is needed about what a combined science option may look like and how it can work as part of a balanced curriculum, with the appropriate amount of time allocated. SCORE is happy to lead this discussion on behalf of the Department for Education and to present a model for science qualifications that accommodates science for progression and science for the citizen.

22. SCORE is very concerned that the EBacc may allow a student to study just two of the three science disciplines. This would be detrimental both to uptake in the sciences beyond 16, and to the education of all, since it would limit their exposure to one or more of the sciences. It will also have implications for progression, since students may not study subjects they subsequently need for moving on to a desired destination (for

example, chemistry is often required for a biology degree). And since it is an accountability measure, there will be little incentive for schools to teach beyond the minimum requirements.

23. It is essential that practical work is assessed more effectively at Key Stage 4 than currently. This requirement should be included in the criteria Department for Education and Ofqual use to appoint a successful bidder. Practical work is intrinsic to science teaching and learning; it helps students to ‘think like a scientist’, develop practical skills and to conceptualise knowledge and understanding⁴. Practical skills are a major component of science at higher levels, yet we are failing to provide a good grounding in both the technical and conceptual skills for progression. SCORE will include a set of practical work assessment objectives for the sciences in its public guidelines for Key Stage 4 qualifications that will be made publicly available in early February 2013 (see paragraph 16).
24. How best to assess practical work is a matter for awarding organisations but we recommend the following are taken into consideration:
- Increasing time constraints created by an overloaded school curriculum have resulted in less time being available for practical activity. Better students are encouraged to study too many subjects.
 - ‘Teaching to the test’, with students drilled in the particular skills needed for a controlled assessment, as well as accountability pressures (including a culture of re-sits) have resulted in less time being available for practical activity.
 - Anecdotal evidence suggests that reduced provision of technical support and resources has resulted in lower quality practical work. Teachers also report that if practical work is not assessed it will not be taught.
 - More evidence is needed to claim practical assessment can be effectively achieved through a written exam. More time should be given outside of this consultation to recommend a suitable way forward. However, SCORE also recognises that there are concerns about the impact of controlled assessment on teaching and learning, and we welcome attempts to address this issue.

Diversity and equal opportunities

25. We note above that the EBacc rules may allow students to study only two of the three sciences; this could have a negative impact on uptake of individual sciences among particular groups. In 2011, only 20% of physics A-level students were girls and there is also evidence that socioeconomic status has an impact on uptake of science post-16, with schools with a higher proportion of free school meals having lower percentages of students going on to physics.⁵ This is likely to be exacerbated by any measure that allows the dropping of one of the sciences.

⁴ see SCORE report, *Practical Work in Science: A Report and Proposal for a Strategic Framework* (2008).
<http://www.score-education.org/media/3668/report.pdf>

⁵ Institute of Physics, *It’s Different for Girls* (2012),
http://www.iop.org/education/teacher/support/girls_physics/file_58196.pdf

26. There is an assumption made in the consultation document that all children with a good education should be able to achieve a level of performance beyond the minimum level currently required to achieve a grade C at GCSE. This only makes sense with a criterion referenced rather than norm referenced assessment system. Whilst we acknowledge that we need to raise students' expectations and aspirations, there is no evidence to demonstrate this can happen within the existing system and we are concerned that policy is being decided based on a false assumption. It would be useful to define what sort of threshold is being assumed by the achievement of a grade C in terms of performance and progression.
27. There is also no evidence to suggest that a student would be any better prepared at 17 or 18 to take the EBC, which could lead to students having to re-take examinations repeatedly. This would be very de-motivating, when it may be more appropriate for them to be pursuing different education or training routes.
28. The EBacc as it is currently described allows for the recognition of only a narrow type of achievement in a particular set of academic subjects at 16, which will reinforce the idea that vocational study has less value than an academic route. It is important that qualifications used for accountability are able to point to a range of different destinations, to avoid disadvantaging a large proportion of students.
29. If EBCs are to play any role in accountability there need to be mechanisms in place to ensure there are no perverse incentives for teachers to advise students against entering for the exam. We do not support the proposed 'statement of achievement' for those not entered for qualifications at 16, unless this statement is given to all students and has currency among the users of education. It risks appearing as a 'statement of failure'.
30. We are also unconvinced that a single tiered assessment system can accurately and fairly measure achievement across all abilities whilst raising aspirations and increasing the motivation of all. The early years of GCSEs existed in a three tiered system (foundation, intermediate and higher) which allowed for a greater degree of stretch and challenge in the assessment items, while also allowing for lower-achieving students to receive a grade demonstrating their level of achievement. Moving from a three-tiered to a two-tiered model has made it harder to differentiate between the most able students because there are not enough questions at the A/A* level. A single tiered system would exacerbate this problem, particularly if one of the purposes of the assessment was to provide certification:
 - It would fail to enable students at either end of the achievement scale to demonstrate their abilities effectively, resulting in question papers that are off-putting for those that find it too difficult and trivial for those that find it too easy. It would thus be unlikely to meet one of the consultation's main aims, which is to provide differentiation for strong performance and recognition of lower levels of performance.
 - Assessments would also have to be very long in order to allow for sufficient items to assess the full range of ability.
31. There are other models of assessment that could be investigated to address this issue, for example a paper to demonstrate competency in a subject taken by all, plus an additional paper taken by high achievers to provide differentiation. SCORE recommends

that the Department for Education consults with assessment experts to explore these possibilities.

Timing of the reform

32. SCORE does not believe the proposed timeframe for reform is achievable. The Department for Education must allow more time to consider these major changes and ensure there is thorough engagement with subject and assessment experts. We recommend the government establishes a group of experts to consider these changes in more depth, to define a purpose, and to define a suitable process for qualification development, taking into account all the unintended consequences. In the past QCDA would have informed Government's thinking but there are still other avenues where advice should be sought, for example much more use could be made of Ofsted and the Standards and Testing Agency.
33. The timetable for reform must allow sufficient time to pilot any new qualifications. It should also be consistent with the timetable for reforming A-levels, while minimising the impact on individual cohorts of students. The Department for Education must also consider how changes will be communicated to employers and other stakeholders
34. While we understand the logic of phasing subjects it will lead to greater confusion among users of qualifications as to the worth of each qualification, and we do not believe strategically important subjects such as the sciences should act as 'pilot subjects' for the reform. With a longer timeframe it would be possible to roll out all subjects at once.

Assessment and grading

35. We see no benefits of having multiple grading systems among awarding organisations and across different qualifications. This will lead to greater confusion among the users of qualifications and we would strongly recommend that the grading system used for all EBCs is also used for all A-level.
36. However, the grading system should provide more information on student performance. For example, supplementary information could be made available on the students' performance in relation to the national entry of a subject via a normal distribution, marks could be given in addition to grades and a detailed breakdown could be provided of students' achievement in assessment components. It will also be important to provide information on the 'exchange rate' between existing and new qualifications, so that employers and other users can make appropriate judgements about candidates.
37. SCORE agrees that the opportunity to take numerous re-sits has had a negative effect on the GCSE qualification: students do not always treat the exam seriously if they know they have the opportunity to re-sit; it offers perverse incentives for teachers to focus on accountability measures; and it can significantly reduce the amount of teaching time.
38. We welcome the move that assessment should limit predictability so that there is no incentive for teachers to tailor teaching specifically to match patterns in the syllabus areas tested or in the types of questions asked. We would also like to see a wider range of assessment methods, including extended writing that would allow students to demonstrate their understanding of more complex scientific processes. However,

39. SCORE has some concerns about the proposed restrictions on examination aids.

Calculators and periodic tables are referred to in the consultation document, but are not examination aids; both are essential for demonstrating mastery of parts of the science curriculum and should be allowed in examinations. There is no intrinsic merit to memorising the periodic table; it is the ability to use the information that it contains that needs to be assessed. Moreover, the difficulty of an assessment is not undermined by allowing the use of these items; in fact, allowing students access to items such as calculators and periodic tables also allows for more complex and demanding assessment that test students' understanding, which would support the Government's aim of further differentiation of able students.