

Education Select Committee Inquiry

How should examinations for 15-19 year olds in England be run?

Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate, a non-teaching department of the University of Cambridge. It is a not-for-profit organisation formed from the examination boards established in the 1850s by the Universities of Oxford and Cambridge, The Royal Society for the encouragement of Arts, Manufactures and Commerce as well as various Midlands-based exam boards.

It was established to raise standards in education and its mission continues to be to promote educational excellence, and to leverage the University's societal impact by reaching out to a broader educational constituency than that which is accessed through the University's core undergraduate teaching and research activities.

The organisation delivers the widest range of qualifications in the world, examines more than 8 million candidates a year and operates in over 160 countries. It employs over 1,800 staff, uses approximately 30,000 examiners and has a turnover of around £270 million, just under half of which is derived from its UK operations.

Additionally, the Group provides a range of educational services. These stretch from advising governments on major educational and curriculum reform programmes to teacher development to helping improve standards of taught English in state primary school systems overseas.

The Group also invests heavily in operational and theoretical research, supported by a total of 63 staff. This supports its core assessment activity, contributes to the development of understanding about the impact of assessment on education and has enabled us to contribute constructively to public discussion and policy debate about educational reform.

In this evidence we address the Committee’s point “*the arguments in favour of and against having a range of awarding bodies for academic and applied qualifications...and the merits of alternative arrangements, such as having one national body or examination boards franchised to offer qualifications in particular subjects or fields*”.

Summary

Typically, concerns focus on four key areas: risks in the system, standards, costs and efficiency. Our evidence shows that a system of multiple awarding bodies offering a range of competitive qualifications - when properly regulated - (A) reduces risk in the system, (B) maintains standards, (C) is cost effective, (D) is efficient, (E) promotes innovation and choice for learners, and (F) generates competitive pressures that ensure constant system modernisation. Covering each of these issues in turn, we have attempted to provide the committee with concrete examples which reflect both past UK experience and best practice abroad. In order that current arrangements might achieve their full potential we have also supplied recommendations for improvement in the current system for the committee to consider.

The evidence from overseas requires careful understanding of the nature of the checks and balances within those systems. Some nations such as Singapore, Hong Kong, France and Finland operate a ‘single board’ model. Conversely, assessment in Queensland, Norway and Sweden is highly devolved. Many of the most successful jurisdictions¹ are of similar size – around 7 million people. This gives a scale which enables high levels of consultation, accountability and engagement – all of which are much weaker in England with its 51 million population.

¹ (as judged by outcomes in TIMSS, PISA and PIRLS)

Meanwhile, other nations are designing systems similar in approach to the UK. The USA has no federal public examinations system and is in the process of developing standards akin to A level criteria. Germany is in the same position, with school assessment operated on a regional state basis and is developing overarching qualification criteria.

A. Risk Reduction

1. Put simply, multiple awarding bodies spread risk.
2. The most error-prone system giving poor outcomes has proved to be the national curriculum assessment tests (SATs). Effectively, a ‘single board’ - a Government department contracting with various agencies for different parts of the process – has delivered serial failure in data collection with severe crises in 1997, 2004 and 2008². The 2008 failure created a crisis of confidence in National Assessment with all 600,000 pupils adversely affected with some schools and young pupils particularly compromised by the failure.
3. In Scotland in 2000 the newly formed single exam board established by the Scottish Executive suffered a systems failure resulting in over 147,000 exam results being rechecked and 120,000 appeals received whilst UCAS warned that 5,000 university places were at risk. Newspaper headlines told of “*Scottish exam chaos*”, “*Exams chief resigns*” and “*Cross words as exam accusations fly*”. The most interesting headline was “*Galbraith accused over exam powers*” as the Scottish Executive sought, when it could not deliver, to distance itself from the machine it had created.
4. By contrast multiple awarding bodies spread risk in the system and avoid single points of failure. For example, while deeply regrettable, the errors in Summer 2011 GCSE and A level question papers afflicted much smaller

² <http://www.parliament.the-stationery-office.co.uk/pa/cm200809/cmselect/cmchilsch/1037/103704.htm>

numbers than would have been the case under a single exam board structure as there is a spread of candidates across all three boards. Contrast this with the problems with a maths paper in this year's French Baccalaureate, where there is a single provider, which caused problems for 165,000 candidates.³

5. GCSEs and GCEs are delivered by different organisations operating under the same regulatory requirements regarding both quality of service and standards of assessment. With a sound regulatory regime, quality and choice is maintained whilst risk is reduced.

B. Standards

6. Creating a single Awarding Body does not of itself secure standards. There remains the issue of the need to ensure standards over time and standards between syllabuses in the same subject. For example, it is almost certainly the case that there will need to exist a range of syllabuses at level 2 and level 3, particularly in Maths⁴. Even if there existed only a single board, or one contract for each subject, a variety of syllabuses would still be needed in order to meet genuine variation in needs. A single board therefore does not eliminate the demanding task of managing comparability between syllabuses in the same subject, at the same level, and between different subjects.

7. The system needs a robust approach to comparability, not administrative re-organisation. Without a sound methodology for comparability, the creation of a single board, of itself, would not help resolve this deeper problem of maintenance of standards and management of equivalences. If there are robust methods in place, the system can handle syllabuses from a variety of bodies. The role of the

³ <http://www.france24.com/en/20110623-france-hunt-math-examination-cheating-education-baccalaureate-chatel>

⁴ Mathematical Needs; Mathematics in the workplace and in Higher Education, Nuffield, June 2011
[http://www.nuffieldfoundation.org/sites/default/files/files/ACME_Theme_A_final%20\(2\).pdf](http://www.nuffieldfoundation.org/sites/default/files/files/ACME_Theme_A_final%20(2).pdf)

regulator should be to use such robust methods continually, regardless of the number of boards.

8. Recommendation: Regulators should expand the number of comparability studies they undertake, using generally accepted methodologies and people generally accepted as experts in comparability, in line with their principal objectives.⁵

9. In fact, the English regulator Ofqual's objectives are to maintain qualification standards, regulate assessment practices and standards, promote public confidence in the system, promote awareness of the qualifications landscape and ensure efficiency and cost effectiveness within the system⁶. Two of these objects deal directly with standards; however Ofqual spends much of its efforts servicing the more diffuse public confidence objective through general media and PR programmes. The objective would be better served through the programmatic delivery of robust comparability studies that investigate whether boards are maintaining standards.

10. We believe the public confidence objective is much too general and diffuse and that it should be dropped, so as to make it clear that Ofqual's principal objective should be the maintenance of standards.

11. Recommendation: that Ofqual's public confidence objective is removed allowing it to deliver public confidence by performing its fundamental task.

⁵128 (1) Ofqual's objectives are—
(a) the qualifications standards objective,
(b) the assessments standards objective,
(c) the public confidence objective,
(d) the awareness objective, and
(e) the efficiency objective.

⁶ <http://comment.ofqual.gov.uk/regulating-for-confidence-in-standards/overview/ofquals-statutory-objectives/>

12. Constant reform, re-structuring and reorganisation of the qualifications system is a major threat to standards. This interference also applies to the reaccreditation cycle for some qualifications. This is where qualifications are only accredited for a limited number of years and then have to be reaccredited, whether or not they need updating.

13. The UK has a long history of serial political interventions. By contrast, in France, education is broadly regarded as a critical structural institution which should endure and should not be prone to politically-motivated change. The checks and balances are subtle, but the separation of day-to-day and strategic policies from the rational strategic management of education is acculturated into the system and national consciousness.

14. Democratic governments have an entirely legitimate interest in control of education and training systems – but ill-founded serial intervention in examinations does nothing to enhance the quality of provision. The point of maximum risk in maintaining standards year on year is when a qualification changes.

15. Recommendation: The frequency and scope of change in qualifications should be reduced.

C. Cost

16. The primary purpose of examinations is to evaluate what students have learnt and been taught. Education expenditure in England was around £70 billion (2008), with some £27 billion being spent on secondary education.⁷ It is estimated that some £625 million (including Ofqual costs) is spent on examinations in England, including many of those outside the scope of this

⁷ http://www.hm-treasury.gov.uk/d/pesa_2010_chapter5.pdf

Inquiry, mostly in secondary education.⁸ Similar ratios exist in the other nations of the United Kingdom. This represents a not unreasonable quality assurance cost of around 2.5%.

17. The economic rationale for merging exam boards rests on the belief that it would achieve significant economies of scale. However, at best the impact would be minimal as the number of candidates, exams and examiners would not decrease, though there would be some reduction in overhead costs. However, it is not evident that this would justify the level of operational risk associated with such consolidation; the historical experience shows that exam board mergers in the 1990s failed to yield significant savings.

D. Efficiency

18. In its terms of reference the Committee explicitly identifies franchising, and the specialisation of exam boards in specific fields, as an alternative to current arrangements. The underlying assumption is that this is a more efficient way of utilising available resources. This section deals specifically with that issue.

19. It is argued that in certain subjects (e.g. maths) there are not enough specialist teachers to deliver sufficient examiners at all levels of experience and expertise to a multiplicity of boards. Therefore, by franchising for individual subjects, all subject experts will move to the winning board which will be in a position to use only the best, allowing the sub-optimal to fall by the wayside. As a result quality and efficiency in examining would improve.

20. However, concentrating expertise in this way would eliminate any possibility of competition in future years, because no other board would be capable of bidding when the contract came up for renewal. This is the situation with the

⁸ Annual Qualifications Market Report, Ofqual, May 2011

Government's Key Stage tests. For some years now, Edexcel has been the only credible bidder for the delivery of the contract.

21. This arrangement is fine only while the Government is happy with its chosen supplier. In 2007 the Government sought to reintroduce competition and international testing company, ETS, was given the contract. In July 2008, around 1.2 million pupils heard that the results for their National Curriculum tests at Key Stage 2 and 3 would be delayed. Considerable blame was placed at the doors of both the Department for Education and its 'arms-length' operator, the QCA. The political storm and public outcry resulted in both the Sutherland Enquiry⁹ and an investigation by this Committee's predecessor¹⁰.

22. While the 'pool' of available examiners in certain subjects is not as deep as all boards might wish, there are alternatives to franchising which do not inhibit competition. Very many fine teachers are prevented from becoming examiners by their senior management who are reluctant to give time off for training, standardisation and administration. To accept that the way to deal with an artificial minimisation of expertise is to focus it in one place seems to us to be placing the cart before the horse.

23. What is needed is for the teacher leadership to encourage those with expertise to engage with awarding bodies at all levels of examining. This not only makes a significant contribution to the health of the whole system but also improves the expertise of the teacher and hence the school. In this way the total number of potential examiners can be expanded, with any number of boards being able to recruit the best.

24. Recommendation: That Head teachers and Principals of colleges be encouraged by government to support teachers who wish to do so to become examiners.

⁹ The Sutherland Enquiry, December 2008 <http://sutherlandinquiry.independent.gov.uk/>

¹⁰ Children Schools and Families Select Committee Report. <http://www.publications.parliament.uk/pa/cm200809/cmselect/cmchilsch/205/20502.htm>

E. Innovation and Choice

25. Innovation can be driven centrally or through competition. Historically the most successful qualifications innovations have emerged from awarding organisations successfully triangulating the needs of government and society with the needs of, on one hand, universities and employers, and on the other, educators and trainers wanting to offer individuals satisfying, stimulating and manageable learning. Examples from 1980 onward include Suffolk, Nuffield and Salters in the Sciences, Ridgeway History, Cambridge History Project, MEI Maths, Critical Thinking, OCR Nationals and CLAIT with integrated Microsoft units. This period saw schools and universities develop 'curriculum projects' to raise achievement and participation and then turn to awarding bodies to supply appropriate, high quality assessment and certification.

26. Innovation and choice go hand in hand. A multiplicity of exam boards enabled these school groups and universities to approach several different boards, some of whom saw the opportunities for different syllabuses, whilst others did not. If only one exam board existed a single negative response would terminate an innovation before it saw the light of day. In contrast, under the present system, the production of a new syllabus which was successful has served to encourage other boards to enhance their syllabuses to meet the clear demand. Such contestability rarely arises where only one producer exists.

27. By contrast centrally derived innovation has given the country Diplomas, GNVQs and NVQs; the first two ultimately failed and the third is subject to a persistent lack of public credibility.

28. Looking abroad, even those countries with highly centralised systems, like France and Singapore, find themselves having to open up their systems to outside providers. They do this to enrich the curriculum, using syllabuses not

otherwise available in the national curriculum, to ensure that the national curriculum is exposed to ideas from abroad, to check the quality of their own examinations and drive improvement. Countries may even go on to design their own version of the examination, as in Colombia,¹¹ but the fact remains that these countries are aware of the need for some form of mechanism help drive innovation, choice and standards.

29. Recommendation:

UK regulators should give force to their general duty to “have regard to the desirability of facilitating innovation in connection with the provision of regulated qualifications”¹² by more flexible application of regulatory requirements for public qualifications.

F. Modernisation

30. Assessment and public qualifications have become technology-intensive, requiring very high levels of investment in e-assessment, logistics, e-marking and so on. This is a direct result of competitive stimulus. Schools and centres demand increasingly detailed, complex information on attainment in order to more accurately match students to programmes. Demand is rising for tests to be taken when the student is ready which requires enhanced computer-based testing (CBT). Exam boards are, rightly, under pressure to improve the quality of marking and grading, best achieved through on-line applications which allow entirely new forms of quality assurance including real-time monitoring of marker performance. Whenever one board achieves an advance, the others compete vigorously to catch up and overtake. A single body, or a franchised body, would be under no such pressure.

¹¹ In Colombia the Ministerio de Educación Nacional launched a National Bilingual Project designed to equip the population with English skills for work and higher study. Cambridge was invited to help with a benchmarking project and to assist in producing a new series of state exams linked to international standards. Following on from the project, the government again requested support to develop new national English language tests, to be incorporated into the Examen de Estado, the national school exit exam in Colombia, taken by more than 500,000 school-leavers in 2007.

¹² <http://www.legislation.gov.uk/ukpga/2009/22/section/129>

G. Improving the current system

31. Different learners learn in different ways and different routes through the education and training system requiring different forms of assessment and certification. With young people compelled to participate in learning programmes from 5-18, it makes sense in personal, economic and social terms to provide programmes which maximize the engagement, motivation and attainment of learners. A rich qualifications catalogue is needed to achieve this, enabling education providers to develop optimized learning programmes.

32. Although it is often argued that there are too many qualifications, the total number of active qualifications (5,500) was comparable to countries such as Germany in 2006/07. Since then, as a direct result of Government policy action designed to decrease the number of qualifications, there has been a massive increase, with Qualifications and Credit Framework (QCF) qualifications increasing from zero to 1,700¹³. This contrasts with the figure commonly used figure of 'around 20,000' qualifications which is a myth.¹⁴

33. A diverse range of qualifications can only happen with multiple, competing, exam boards innovating and working with schools, HE and employers. However, the UK system is not operating optimally. Although Ofqual originally set out to regulate the overall system for delivering qualifications, it has re-introduced the micromanagement of 'public qualifications' in addition to high-level regulation, thus increasing the regulatory burden and stifling innovation.

¹³ <http://www.ofqual.gov.uk/standards/92-articles/594-available-qualifications>

¹⁴ This figure derives from a request by Sir William Stubbs (then Chief Executive of FEFC) to the FEFC data unit, prior to the formation of QCA in 1997. It was simply the total number of qualifications known to FEFC. Crucially, it did not exclude duplicate entries on the database (errors); the same qualification offered by a different awarding body; old versions of revised qualifications; and completely redundant qualifications.

34. Recommendation: Regulators should focus on high-level definitions of needs and be flexible in accrediting assessments that reach the same goal using alternative methodologies.

35. The efforts of the regulator, like those of the Committee, have concentrated on what might called 'high stakes' assessments i.e. those giving access to the upper levels of education or employment produced by five UK exam boards.¹⁵. However, there are in fact 160+ registered awarding bodies which in turn offer a multitude of assessed (most vocational) qualifications. These in turn are viewed as 'low stakes' qualifications.

36. We believe a differentiation of 'High' and 'Low' stakes is fundamentally wrong. All exams impact upon those taking them. Cambridge Assessment research indicates that seemingly 'low stakes' assessment can have a powerful effect on learners' self-esteem and dramatically affect their decisions about the things at which they are 'good' and 'bad'. In turn, this can affect their decisions about which directions to take, altering the life chances of a candidate sometimes by a greater factor than a 'gifted and talented' candidate failing to get the grade they wanted at GCSE.

37. Ofqual's concentration on 'High' stake qualifications effectively prevents it from being a rigorous regulator in relation to the protection of candidates taking 'Low' stakes qualifications.

38. Recommendation: UK regulators must not narrow their regulatory focus to 'High stakes' qualifications only.

39. The interrelationship of England's qualifications system with those of Wales, Scotland and Northern Ireland has never been seriously addressed. It is notable,

¹⁵ <http://www.ofqual.gov.uk/files/2010-10-19-4776-from-transition-to-transformation.pdf>

for instance, that arguments for centralising are seldom applied to the exam boards under the control of the administrations in Wales, Scotland and Northern Ireland but only to the three 'English' boards.

40. More importantly, similarly titled qualifications (e.g. GCSEs) offered in any country of the UK must be of an equivalent standard if learners are to enjoy employment and higher education mobility throughout. There are already serious doubts as to whether English GCSE standards are diverging. It may no longer be tenable for the English regulator to accept by proxy decisions made by the other UK regulators and vice versa.

41. Recommendation: Serious consideration should be given to the arrangements for ensuring UK-wide comparability of standards, starting with a new consultation on the issue and the appointment of an intra-country comparability 'czar'.