

Westminster Education Forum Keynote Seminar: Reforming England's examination system - quality of marking, market regulation and tackling malpractice

Panel presentation: Competition in the exams market: how the comparability framework works against the essential purposes of assessment and what we can do about it

James Croft, Director, Centre for Market Reform of Education

I want to focus my remarks in this brief presentation on the competition/quality issue to revisit the thesis that supplied the justification for the system overhaul we are now in the midst of. First to ask whether competition has indeed been responsible for so-called dumbing down of qualifications (perhaps better termed 'grade drift'), and then to consider whether with government reforms to re-focus the curriculum and learning on fewer, more challenging subjects, and to introduce a greater level of distinction between the top grades, we can expect a better functioning market to emerge? Well, to address the central question, I think the straightforward answer, despite much bluster by the Education Secretary and in the press to the contrary, is no. Competition has been mistaken for comparability, which is the real issue. It is the effect of poor performance metrics and of the comparability framework, not competition, which has prevented individual boards from introducing more stretching content to exams; investing in the development of more diverse and at the same time more challenging alternative qualifications; and inhibited them becoming the autonomous, self-improving, self-certifying authorities that they might be.

So what is the comparability framework and why is it so unhelpful to meaningful assessment?

The purpose of assessment is to provide information about the level of a young person's understanding, aptitude and ability. The purpose of the comparability framework is to ensure and evidence that all young people have been given an equal opportunity to succeed. The latter is a political construct, the pressure of which, as it has developed, has obscured and undermined assessment's essential purpose, in placing considerations of fair access ahead of their effectiveness as signalling mechanisms for end-users.

This is a difficult subject. It's not that we shouldn't be making every effort to improve education, or that we shouldn't be prioritising resources to those that need them most. Rather the question is whether in trying to equalise, we in fact merely homogenise, and do young people a disservice in the process – and what can be done about this. The fundamental problem with making equal opportunity the overriding goal of the system is that it is very difficult to judge when an opportunity has been extended in such a way that all might have taken advantage of it ... without reference to outcomes ... which is where it all starts to unravel. So, over getting on for three decades politicians have increasingly fixated on GCSE and A level performance and especially higher education (and now Russell Group, or even Oxbridge!) access as the measure of the success or failure of this system as a whole. Perhaps it's time to start asking whether aiming for diverse quality outcomes, entailing progression into a broader range of different training and workplace settings, befitting individuals' different strengths and aptitudes is not a more sensible goal.

For this we need more competition, not less, but competition on a different basis to that we have now. Examining the conditions under which boards must compete in a national qualifications framework should be our priority.

The problems with comparability in respect of standards have manifested themselves in a number of ways. In respect of inter-board standards, the measurement framework disincentivises individual boards' efforts to make assessments more demanding. Statistical screening techniques are used to identify unacceptable levels of deviance between predicted and actual levels of attainment so that grade boundaries may be adjusted to compensate and bring wandering awarding bodies into line. In effect this means that the boards do not own their brands in respect of quality assurance. These are

subordinate to those of the national qualifications – GCSE, A level, etc. – they award. They are therefore not incentivised to compete on this basis.

With inter-subject and assessment standards, efforts to stretch the national brand to accommodate diverse subjects and means of assessment evidently work against what qualifications are designed to do in respect of their signalling function. Instead of allowing boards to develop syllabi for students to explore and discover their aptitude for, and assessments that are appropriate to this end, the qualifications trail for which learning journey might provide the kind of nuanced picture of young people's capabilities we surely want, the requirements of comparability homogenise their learning experience, and actually end up narrowing many students' options.

With comparability of assessment across the full array of vocational and academic-oriented qualifications, it leads to the creation of artificial and misleading frameworks of equivalence. The point is expressly to equalise difference, by means of assigning levels – an inevitably arbitrary exercise, which creates confusion between formal and substantial equivalence, and blurs the distinctions between what different kinds of qualifications are designed to do (a problem recognised and addressed in the Wolf Review, but not ultimately resolved).

Finally, in respect of the question of the comparability of overall levels of attainment over time, (expressly, the long-term) and whether such can be said to have improved over time as a result of government policies. Though the subject of much debate, this is really irrelevant to the question of assessing whether qualifications are performing the function they are designed for. Changes in the wider social, economic, educational, and not least the policy context, and the changes to content that result, make comparisons with 'what it were like in my day' a fruitless exercise. These include a whole plethora of factors, but we lack a methodology, and therefore evidence for which may have had actual, let alone significant bearing on improvements in grades.

Curiously, not unlike the consensus that has developed around the 'competition depresses standards' thesis, policymakers seem to have reached broad agreement over the centrality of Ofqual's role in sorting all this out. The government asked Ofqual to ditch past efforts to provide the reassuring narrative necessary to maintaining faith that improving results meant improving education and step up to challenge both politicians and exam boards through the employment of a tightened up comparability framework. The regulator has stepped up to this and now polices the limits of comparability. The most recent squeeze has been on the many GCSEs in 'peripheral subjects' deemed to be not as challenging as core academic subjects included in the new EBacc performance measure.

This and other government policy-led reforms in this area to date essentially amount to reducing the extent to which these comparability controls need to be applied by homogenising provision. Proposals for franchising by subject, or the nationalisation of provision, would take us even further down this road. It's one answer, but not one that eradicates the comparability problem (between different subjects, assessments, and types of qualification, for example), but one that is not without potentially very unpalatable consequences – neglect of areas of the curriculum and indeed extra-curricular learning, soft skill and character development, and polarised outcomes to mention but a few of the risks. At which point all efforts to tighten comparability to ensure fairer access and outcomes will seem futile.

A better way, which might better reconcile the competing demands of extending opportunity and ensuring rigour, and support the development of the kind of well-rounded educational experience that the education secretary wants to see in his new independent schools, would be to open up a market for information provision about the quality and utility of qualifications. Such a market would inform users of their currency for progression into, in and beyond, various higher educational and training contexts, and enable judgements about their relative value on a range of different outcomes-focused indicators. A market in information provision about the quality and utility of qualifications would encourage greater responsiveness to complex, nuanced and ever-changing end-

user requirements too – far more effectively, at any rate, than government or regulator initiated consultations are able to.

Exam boards should succeed or fail on the basis of the quality of their qualifications and their service alone, not on the basis of whether they can stay the right side of politics. There are four things they need to get right that are relevant to the interests of end-users to ensure the utility of their qualifications – each of which need to be considered in relation to the profile of learner for whom they are designed. These relate to the content (specifically its relevance and real-world utility), the level of challenge or difficulty, what grades mean (how well a student has done in relation to certain criteria or the performance of their peers), and the reliability and validity of assessments (whether they work as tests and measure what they claim to measure). None of these require the policing of a regulator: exam boards need to attend to these issues anyway to remain competitive, and indeed do so in the context of international markets where they are developing new independent and evidently valuable qualifications all the time. These don't need rubber-stamping and fitting into artificial regulatory frameworks to maintain their credibility.