

Written statement from Chair of Ofqual to the Education Select Committee

On behalf of my Board, I welcome the opportunity to give evidence to the Select Committee and provide, by way of a written statement, some introductory comments.

Above all else, we want to make clear that we are sorry for what happened this summer: the distress and anxiety it has caused for many students and their parents; the problems it has created for teachers; and the impact it has had on higher and further education providers.

In March, Ofqual was consulted by the Secretary of State on how to manage school qualifications in the context of a pandemic. Our advice at that time was that the best option in terms of valid qualifications would be to hold exams in a socially distanced manner. We also set out alternative options including the use of standardised teacher assessments and the risks associated with them.

On March 18, the Secretary of State for Education took the decision to cancel exams this summer. The loss of schooling and the likely parental concerns about sending children back into schools to take exams meant that exams were not considered a viable option.

We were asked to implement a system of grading using standardised teacher assessments, and directed to ensure that any model did not lead to excessive grade inflation compared with last year's results. The primary objective was to allow young people to progress with their lives, whether to sixth form, college, university, work or training. Given that they could not demonstrate their abilities in summer exams, our approach was supplemented by an opportunity to sit exams in the autumn.

The principle of moderating teacher grades was accepted as a sound one, and indeed the relevant regulatory and examination bodies across the four nations of the United Kingdom separately put in place plans to do this. All the evidence shows that teachers vary considerably in the generosity of their grading – as every school pupil knows. Also, using teacher assessment alone might exacerbate socio-economic disadvantage. Using statistics to iron out these differences and ensure consistency looked, in principle, to be a good idea. That is why in our consultations and stakeholder discussions all the teaching unions supported the approach we adopted. Indeed when we consulted on it, 89% of respondents agreed or strongly agreed with our proposed aims for the statistical standardisation approach.

We knew, however, that there would be specific issues associated with this approach. In particular, statistical standardisation of this kind will inevitably result in a very small proportion of quite anomalous results that would need to be corrected by applying human judgment through an appeals process.

For example, we were concerned about bright students in historically low attaining schools. We identified that approximately 0.2% of young peoples' grades were affected by this but that it was not possible to determine in advance which cases warranted a change to grades. That is why the appeals process we designed and refined was so important. But we recognise that young people receiving these results experienced significant distress and that this caused people to question the process.

The statistical standardisation process was not biased – we did the analyses to check and found there was no widening of the attainment gap. We have published this analysis. Indeed, 'A' and 'A*' grade students in more disadvantaged areas did relatively better with standardised results than when results were not standardised.

However, the impossibility of standardising very small classes meant that some subjects and some centres could not be standardised, and so saw higher grades on average than would have been expected if it had been possible to standardise their results. This benefitted smaller schools and disadvantaged larger schools and colleges. It affected private schools in particular, as well as some smaller maintained schools and colleges, special schools, pupil referral units, hospital schools and similar institutions. We knew about this, but were unable to find a solution to this problem. However, we still regarded standardisation as preferable because overall it reduced the relative advantage of private schools compared to others.

Ultimately, however, the approach failed to win public confidence, even in circumstances where it was operating exactly as we had intended it to. While sound in principle, candidates who had reasonable expectations of achieving a grade were not willing to accept that they had been selected on the basis of teacher rankings and statistical predictions to receive a lower grade. To be told that you cannot progress as you wanted because you have been awarded a lower grade in this way was unacceptable and so the approach had to be withdrawn. We apologise for this. It caused distress to young people, problems for teachers, disrupted university admissions and left young people with qualifications in which confidence has been shaken. It will affect those taking qualifications next year who are competing for the same opportunities as those who received this year's grades.

We fully accept our share of responsibility in this. Throughout the whole period we worked in close partnership and transparently with the Department for Education. We also consulted widely including with exam boards and with relevant education unions to ensure the proposals had their support.

There has been much discussion about the design of the algorithm. Many designs were considered and many proposals put forward. The suggestion has been made that a different model might have led to a different outcome. But the evidence from this summer, including from similar models implemented and withdrawn in Scotland, Wales and Northern Ireland indicates a much more fundamental problem. With hindsight it appears unlikely that we could ever have delivered this policy successfully.

What became apparent in the days after issuing A level results was that

neither the equalities analyses, nor the prospect of appeals, nor the opportunity to take exams in the autumn, could make up for the feeling of unfairness that a student had when given a grade other than what they and their teachers believed they were capable of, without having had the chance to sit the exam.

Understandably, there is now a desire to attribute blame. The decision to use a system of statistical standardised teacher assessments was taken by the Secretary of State and issued as a direction to Ofqual. Ofqual could have rejected this, but we decided that this was in the best interests of students, so that they could progress to their next stage of education, training or work.

The implementation of that approach was entirely down to Ofqual. However, given the exceptional nature of this year, we worked in a much more collaborative way than we would in a normal year, sharing detailed information with partners.

We kept the Department for Education fully informed about the work we were doing and the approach we intended to take to qualifications, the risks and impact on results as they emerged. However, we are ultimately responsible for the decisions that fall to us as the regulator.

We believe it is important that we do not leap to inaccurate conclusions prematurely. It will take time to fully understand everything that happened here, less than three weeks after results day. But there are already some important lessons to be learned from this summer:

- any awarding process that does not give the individual the ability to affect their fate by demonstrating their skills and knowledge in a fair test will not command and retain public confidence
- the original policy was adopted on the basis that the autumn series would give young people who were disappointed with their results, the opportunity to sit an examination. However, the extended lockdown of schools and the failure to ensure that such candidates could still take their places at university meant that this option was, for many, effectively removed. This significantly shifted the public acceptability of awarding standardised grades
- it is easy for people to believe that a policy is fair at the overall level, but this belief changes very quickly when the impact is felt at an individual level. It is not clear to us that a more effective communications effort would have overcome this, but to be successful it would have to have engaged multiple levels of communication, not simply the activities of the regulator
- a 'better' algorithm would not have made the outcomes significantly more acceptable. The inherent limitations of the data and the nature of the

process were what made it unacceptable

The blame lies with us collectively – all of us who failed to design a mechanism for awarding grades that was acceptable to the public and met the Secretary of State's policy intent of ensuring grades were awarded in a way consistent with the previous year.

To try to deliver comparable qualification results in the absence of students having taken any assessments (examinations) proved to be an impossible task. It is now our collective responsibility to learn the lessons and to establish a way forward that can command public confidence and give students what they need to progress, even in difficult circumstances.

Roger Taylor

Chair, Ofqual Board