ofqual

Executive summary

Awarding GCSE, AS, A level, advanced extension awards and extended project qualifications in summer 2020: interim report

The 2020 exam series has been exceptional. Early in 2020, when coronavirus (COVID-19) was first reported, we modified our normal contingency planning with exam boards to include widespread staff absences and/or large numbers of students being unable to take one or more of their exams due to illness or self-isolation.

On 18 March 2020, the Secretary of State announced[1] that the summer 2020 exam series would be cancelled in order to help fight the spread of coronavirus (COVID-19) and that students due to sit the exams would be awarded a grade based on an assessment of the grade they would have been most likely to achieve had exams gone ahead. On 23 March, in a written statement to Parliament,[2] the Secretary of State explained the government's intention that results would be issued to this year's cohort based on a range of evidence and data, including performance on mock exams and non-exam assessment. This would be achieved by exam boards producing calculated grades ensuring the distribution of grades follows a similar pattern to that in other years, so that this year's students do not face a systemic disadvantage as a consequence of circumstances this year.

Since then, we have been working with exam boards to enable the award of grades for GCSE, AS, A level, Extended Project Qualification and Advanced Extension Award this summer, so that students can move on to sixth form, college, higher education, training, apprenticeships or employment.

Our aim in this work was to use an approach that was, as far as possible, fair to students who had been unable to sit their exams this summer. That approach should also ensure that the results issued had a similar value to grades issued in any other year, so that those using them to select students (sixth forms, universities, employers, etc.) could have confidence that their worth was in line with previous years. A critical factor in achieving that was maintaining overall national standards relative to previous years.

We considered many different options, but it was apparent that the best judges of the relative ability of students in a school or college were the teachers who had been

[1] Schools, colleges and early years settings to close.
[2] Impact of Covid-19 on Summer Exams: Written statement - HLWS170

preparing these students for their exams, tracking their progress relative to target grades, and, in the case of A level students applying to higher education, providing estimated grades.

We therefore asked teachers to provide, for each student for each subject they were entered for, a centre assessment grade (CAG) which represented the grade that student would have been most likely to achieve if teaching and learning had continued and students had taken their exams as planned.

We also asked teachers to provide a rank order of students for each grade for each subject. There were several reasons for this. First, we know from research evidence that people are better at making relative judgements than absolute judgements and that teachers' judgements tend to be more accurate when they are ranking students rather than estimating their future attainment. The research literature suggests that, in estimating the grades students are likely to achieve, teachers tend to be optimistic (although not in all cases). That is not surprising, teachers want to do the best for their students, and the analysis we carried out immediately after CAGs were submitted bears this out.

Exam boards adapted their IT portals to enable schools and colleges to submit over 5 million CAGs and rank order positions in early June 2020. This was achieved despite teachers often working remotely with limited access to school premises, and with exam board technical teams also working remotely.

Our initial analysis of the CAGs showed that they were, in general, optimistic (although not always) and the combined effect would be likely to lead to overall national results that were implausibly high. If we had awarded grades based on CAGs we would have seen overall results increase by far more than we have ever seen in a single year. At A level[3], we would have seen the percentage of A* grades go up by 6 percentage points from 7.7% of grades in 2019 to 13.9% of grades this year, and the percentage of grades that were B and above increase by over 13 percentage points from 51.1% in 2019 to 65% this year.

This optimism was not surprising. It is what is suggested in the research literature and data published every year by UCAS – schools and colleges tend to be optimistic when estimating the grades that students are likely to achieve. Our interviews with teachers, after CAGs had been submitted, confirmed this trend. Almost all the teachers we interviewed told us that they had generally predicted how the students would perform on a 'good day'. Although they knew that every year some students

[3] Note that 2019 figures are for England only as reported by JCQ.

underperform or have a bad day, this was not the basis of their judgements. This might be as expected, but the cumulative effect of this optimism, if reflected in the final results, would have undermined confidence in those results.

Standardisation was not solely implemented to ensure that grades were not, overall, excessively high this year. The key purpose was to ensure fairness to students within the 2020 cohort. Without standardisation there was the potential for students to be unfairly advantaged or disadvantaged, depending on the school or college they attended and the approach they took. A key motivation for the design of the approach to standardisation that we took was to remove this potential inequality and, as far as possible, ensure that a grade represents the same standard, irrespective of the school or college they attended.

Given the circumstances this summer, we had no opportunity to put in place a system of national standardisation, to guide teachers in making their judgements consistently across the country. It was, though, essential that we put in place a mechanism to standardise those judgements being made in many thousands of different schools and colleges, in the interests of fairness to students. It was likely that different schools and colleges would take different approaches to generating CAGs and rank orders and that was likely to generate different levels of optimism in different centres (or perhaps in different subjects). That is what we saw.

It was important, then, to have a system to standardise teachers' judgements across schools and colleges, and that all exam boards took the same approach. A parallel can be drawn here with moderation of teachers' marking of coursework (non-exam assessment) in a normal year. Even when exam boards provide marking criteria, training materials and events to guide teachers in their marking, exam boards moderate the marking to ensure a common standard is applied, in the interests of fairness to students overall.

The rank order data allowed us to make fine-grained adjustments which is fairest for students. For example, with this information we could prioritise students for an upwards or downwards adjustment where necessary. Without the facility to perform this fine-grained adjustment it may have been necessary to adjust larger groups of students leading to an over or under adjustment.

In April and May we worked with technical experts across the sector to test a range of different statistical standardisation models using data from previous years. In selecting the final model, we chose the one that most accurately predicted students' grades in a way that did not systematically affect groups of students with particular protected characteristics. We also considered operational issues – how easy it was to implement the approaches consistently across all four exam boards – and

transparency – how easy it was to explain to schools and colleges how the model worked.

Our preferred model – known as the Direct Centre Performance model (DCP) – works by predicting the distribution of grades for each individual school or college. That prediction is based on the historical performance of the school or college in that subject taking into account any changes in the prior attainment of candidates entering this year compared to previous years. This was fine-tuned to take account of known issues such as centres with small cohorts of students, small-entry subjects, and tiered subjects. Decisions were also made on the number of years of historical data included in the model. The details of these decisions are set out in this report and are formalised in the regulations[4] we put in place for summer 2020.

Where schools and colleges had a relatively small cohort for a subject – fewer than 15 students when looking across the current entry and the historical data – the standardisation model put more weight on the CAGs. Since small teaching groups are more common for AS and A level than for GCSE, and given that the CAGs tended to be optimistic, it means that the outcomes in some AS and A level subjects are much higher this year. However, there is no statistical model that can reliably predict grades for particularly small groups of students. We have therefore used the most reliable evidence available, which is the CAGs.

Overall, A level results in England have increased by 2.4% at grade A and above compared to 2019. This is a larger change than observed in a typical year (for example, there was a 1% decrease in outcomes between 2018 and 2019).

Across all subjects and all centres, 96.4% of final calculated grades are the same as, or within one grade of the CAG submitted. A small percentage were adjusted by 2 grades or more, in some cases because it appeared that the centre's CAGs were very much higher than the historical results in the centre.

In any year, there is measurement uncertainty in the assessment process. This may be due to exam questions being asked that match well with one student's strengths but are poorly matched with the strengths of another student with the same overall level of ability. This year, there is also uncertainty in the results and the challenge is heightened by the absence of any formal assessment information on which to assess students. Based on the testing of the approaches applied this summer using results data from 2019, 51 of the 55 A levels tested had accurate predictions for more than 90% of students within plus or minus one grade[5]. This figure was lower for

[4] [Requirements for the calculation of results in summer 2020](#)
[5] These A level figures are calculated once entries from schools and colleges with fewer than 10 entries in the subject have been removed

GCSE (12 out of 22 subjects) which was likely due to a combination of the grade scaling being longer for GCSE compared to A level and some limitations of the testing. GCSE English language, English literature and mathematics all had above 99% of students receiving results accurate within plus or minus one grade. Overall, the levels of predictive accuracy are broadly comparable to measures of marking consistency across an equivalent range of subjects.

To understand the impact of potential advantage or disadvantage across different demographic and socio-economic groups we have also performed an equalities analysis of calculated grades. The analyses show no evidence that this year's process of awarding grades has introduced bias.

This interim report provides a description of the process for all qualifications and presents analyses of CAGs and calculated grades for AS and A level. Analysis of GCSE CAGs and calculated grades will be published on GCSE results day. A final report will be published later in the year when we have completed our evaluation of this summer's results.

Throughout the development and testing of the model, and in its implementation, we have taken all possible steps to ensure the process is as fair as it can be and, where possible, have taken design decisions in the students' favour. For example, we have used calculations that assume that all students would have attended for all of their assessments this summer. In reality this would not have happened.

We know that, just as in any year, some students will be disappointed with their results. Some students may think that, had they taken their exams, they would have achieved higher grades. We will never know. But for those students who do wish to improve their grades, there will be an autumn exam series.

Where possible, we have urged sixth forms, colleges and universities to be flexible in their selection this year. Overall, the results delivered this year will have met the aim of enabling large numbers of students to move on to the next stages of their lives.