Exams: round one

The run-up to examinations in 2020

In the run-up to the announcement that the 2020 exams round would be cancelled, DfE had started to consult with England's exams regulator, Ofqual, and a limited number of others on what the consequences of this might be. This was part of the "lots of work" that Jonathan Slater had obliquely informed the Public Accounts Committee about.

No one, from the prime minister down, wanted to cancel exams. But as one DfE insider put it, such was the uncertainty over for how long schools would be closed and how the pandemic would develop that "we could just not get assurances that it would be safe to have young people in schools sitting exams in the summer. So the decision was to go with certainty, even if it was far from ideal." The department put up "all the options you would expect from the civil service", including whether there should be 'school leaving certificates' rather than grades.

One of those closely involved in the consultation said:

> **"There was definitely a political concern around what would be publicly acceptable. There was a sense that even if you could do exams in a socially distanced manner, the idea that you were going to say to parents in the middle of a pandemic, 'now please bring your kids into school, we are going to make them sit an exam, even though there might be a health risk – because you can't reduce it to zero' – that seemed like a recipe for disaster. The idea that people might be outraged at having been given a grade that they were unable to control or influence, one that was just generated, was, at the time, much less vivid. And there was a sense that we are in a war-time atmosphere. If things were not perfect, people would understand."**

Ofqual's initial stance was that ideally exams should be held in a socially distanced manner. Failing that, they should be delayed, and failing that it should look at "some form of calculated grade" – or even simply issue teachers' certificates for A-levels, combining that with a big expansion of university places.[43] In hindsight, Roger Taylor, the chair of Ofqual at the time, says that might have been the best solution. Civil servants say it was considered, although Taylor says that to his knowledge "never seriously" – even though, as he points out, in practice, "by a painful, chaotic and unplanned route, that is where all four countries [of the UK] ended up".[44]

On 31 March, without any further consultation with Ofqual, according to Taylor, Williamson confirmed that exams would be cancelled and that pupils would be given calculated results.[45]

INQ000231401_0018

One civil servant says that failing to find a way of holding exams even in a modified form "might be seen as a failure of imagination on our part". Indeed, at the Institute for Government roundtable there was a minority view that some form of examination might indeed have been possible given that infection rates had in fact fallen significantly by the exam season (usually beginning in June). Some other European countries, Germany included, did manage to hold them – although quite where each country was at each stage of the pandemic varied. But the majority view that we heard was that cancellation was inevitable.

The 'algorithm'

On 3 April, Ofqual issued an outline guide as to how grades would be awarded.[46] There were queries and concerns, but there was appreciable initial support for the approach.[47] One or two voices even suggested that it might prove fairer than the usual examination system.[48] Grades would be decided by a mixture of teacher assessment and adjustment, with teachers ranking children within their class, with those results then put through "a process of standardisation" – what became known as 'the algorithm'. It would take into account "the expected national outcomes for this year's students, the prior attainment of students at each school and college (at cohort, not individual level), and the results of the school or college in recent years".[49] Notably, England was not alone within the UK in deciding to use an algorithm to adjust results. Scotland, Wales and Northern Ireland all did the same, although the precise nature of the formulae they used differed.

Assessing pupils' performance was not helped by the fact that moderated coursework had come to play a much smaller part of GCSE and A-level awards for many subjects. Given schools were closed in March, only some of it would have been available. But it might have helped. In the years leading up to 2020, results had come increasingly to depend only on the examination, with much less if any coursework contributing to the outcome* (although that is not a view either for against the change).

One absolutely key issue, as we heard from both Ofqual and the department, was that ministers were desperate to avoid grade inflation. It was "a totemic" issue. As one official it:

> **"One of the things that the government has been proudest of since 2010, and one to which they are most committed in respect of school standards, was the work they had done to tackle what they saw as grade inflation under the previous regime. It was a first order principle that was baked in from the beginning."**

Hence the need to standardise results, given that teacher assessment is known to vary with an understandable tendency to over-mark. Others point out that the results needed to be seen to be fair between different schools and colleges and between successive cohorts of children.

* Other than in practical subjects such as art and music.

INQ000231401_0019

Williamson's direction to Ofqual at the end of March reflected the concern over grade inflation: "Ofqual should ensure, as far as is possible, that qualification standards are maintained and the distribution of grades follows a similar profile to that in previous years" was the instruction.[50]

In April, Ofqual issued a fuller consultation on how results would be handled. The consultation debated the weight that should be given to teacher assessments as they were standardised, against putting more weight in the algorithm on the past performance of schools and colleges – the latter being Ofqual's preferred option. In May, it confirmed that the latter would be its approach.[51] Just over half of respondents (54%) supported that in the consultation, with 33% opposed. Once the full details of the algorithm emerged, however, concerns started to grow.

There were worries, for example, that taking a school's past performance into account would disadvantage pupils in improving schools, or with particularly good cohorts (and all schools know that they have "good" years and less good ones). Ofqual's algorithm, it was argued, was incapable of adjusting grades at the individual pupil level, something that Ofqual itself acknowledged at the time, stating that "it is true that a highly statistical standardisation model will operate at centre not individual student level. It cannot reflect the possibility of individual students doing better or worse than would be predicted by the statistics."[52]

As time went by, these concerns became more public.[53] Sir Jon Coles, chief executive of the United Learning MAT and a former schools director within DfE, warned: "We are experiencing a slow-motion car crash towards the least just results season in history."[54] When people complained that improving schools would be disadvantaged, as would very bright pupils in poorly performing schools, Ofqual's response was:

> **"While we recognise that a small number of centres would like standardised grades to reflect their recent or expected improvements in results, our research into GCSE grading shows the performance of centres rarely improves (or deteriorates) consistently in the short term. Because of this, we do not believe it is possible to design a statistically robust model which could reliably predict such improvements or deteriorations."**[55]

Its view was that schools' and pupils' objections to their results would be dealt with by appeals. As early as May, however, Ofqual had highlighted to ministers "the risk of widespread dissatisfaction" with the grades that would be awarded.[56] One vastly experienced school leader said:

> **"I had an extraordinary three months of trying to talk to people in Ofqual, in DfE, and ultimately with ministers about the inevitable problems. And the extraordinary feature of those conversations was that at almost every level of Ofqual and at almost every level in the department, when I explained what the problem was, they seemed to agree it was the problem – and that included ministers ultimately. And yet nobody changed anything."**

According to DfE insiders, ministers were fully aware that there would be a row.

> "One could have had a different debate, and allowed for some grade inflation," one said. "But preventing or at least restricting grade inflation was totemic. It did not need clever civil servants to point out to the secretary of state that when faced, on the results day, with young people disappointed that they had not got what they thought they would get – what their teachers thought on the one hand and what the algorithm had produced – he was not going to win that argument on the news. He had worked it out for himself. He did not think, none of us thought, that it was impossible to sustain the position, despite the onslaught of attack there would be.

> "We thought we were in extraordinary times. But with hindsight that was a poor decision. Up to the top of government, politicians understood the hit they would be taking. Though not, obviously, as bad as it was."

A No.10 insider confirms that. "There were assurances from DfE that it was fine, and that it was the right approach. There would be a row, but it was rideable. That's what No.10 was being told." In practice it became not so easy to ride as teachers began to realise the effects the algorithm would have on individual pupils.

Scotland U-turns on its algorithm

And then there was Scotland. It had used a different algorithm to adjust its exam results. But when they came out on 4 August, 124,000 teacher assessments proved to have been downgraded – a quarter of the total. Some appeared completely inexplicable, down from a predicted A to a C or even lower. University places were being lost as a result. Within a week, on 11 August, Scotland changed course, dumping its algorithm in favour of teacher-only assessments.[57]

On the day Scotland U-turned, Williamson said: "I'm confident the system Ofqual has put in place is fundamentally a fair one that will award the vast majority of students a calculated grade that genuinely reflects the grade they would have achieved."[58] This was a message he kept repeating, even while officials in DfE were growing increasingly worried about "outlier" results (big and seemingly incomprehensible differences between teacher-assessed and calculated grades),[59] something Ofqual had again briefed them about in the week prior to the Scotland change.

Sensing trouble ahead, the day before A-level results for England were published, Williamson promised a "triple lock" through which pupils could take their calculated grade, appeal to receive a valid mock result (without any definition, or indeed any clear idea, of what a 'valid' mock would be in these circumstances[60]) or sit autumn exams.[61]

The day of publication, 13 August, returned an all-time high of A and A*s. But almost 40% of predicted grades – a much higher percentage than in Scotland – were downgraded, with 3.5% going down two grades or more. The media was full of tearful tales from distraught pupils who had lost university places. Williamson urged schools to appeal. Asked on Sky News whether he would give "a cast-iron guarantee" that he would not be forced into the same U-turn as Scotland, he replied: "Absolutely."

INQ000231401_0021

But given that Scotland had not been able to hold the line with a quarter of results downgraded, the idea that England would be able to do so with approaching half of all results was for the birds. As the A-level results came out it was clear that the GCSE ones would be affected in the same way.

On Saturday 15 August, Williamson told *The Times* that there would be "no U-turn, no change" and "this is it". Following Scotland and using predicted grades would produce "rampant grade inflation", he said.[62]

On Monday 17 August, however, the inevitable happened. Pupils were to be given the higher of either their predicted or their calculated grade, with the same applying to GCSEs. Williamson appeared to blame Ofqual for the debacle, claiming he had only become aware "over the Saturday and Sunday" of the scale of the problem, and that he had "consistently asked a large number of challenging questions about the system. Its robustness and its fairness. We'd been constantly reassured about that."

In a round of media interviews the following day, he went further saying that Ofqual "didn't deliver" the fair system expected.[63] It took an excoriating draft resignation letter from Roger Taylor, Ofqual's chair, pointing out just who was responsible for which decision, and a lengthy meeting with Williamson, for DfE then to declare that "we have full confidence in Ofqual and its leadership in their role as independent regulator".[64] By contrast, when Scotland had reverted to teacher-assessed grades, Nicola Sturgeon absolved Ofqual's equivalent body, the Scottish Qualifications Authority, of blame, saying "we did not get this right" and adding that "I don't attach any blame to the SQA for this. This is ministers' responsibility."[65] The contrast could not have been starker.

A week after the English grades assessment U-turn, Sally Collier resigned as chief executive of Ofqual. The next day Jonathan Slater, whose five-year term as permanent secretary at DfE was due to end in May 2021, stepped down. The prime minister declared that it was time for "fresh official leadership". A "Downing Street source" reported that the decision had been made a month prior, claiming it was unrelated to the exams row,[66] and that Johnson had "concluded a while ago that he was not confident in Slater". It was clear that he had been sacked.[67]

Dave Penman, general secretary of the senior civil servants' union, the FDA, accused the government of "throwing him under a bus", declaring that "ministerial accountability is dead". Calls for the secretary of state to resign were not universal, but they were by now legion, including from the Institute for Government,[68] with even senior Conservatives declining to back him.[69]

INQ000231401_0022