Written evidence submitted by Mr Mark Fretten

East Sussex, August 2020

FAO: Select Committee, The impact of COVID-19 on education and children's services.

I would like to comment to the committee about GCSE grades as currently issued on 20th August and the impact an uncorrected oversight on the part of exams regulator OFQUAL is having on affected students. I am a parent who has followed events of the past few months closely and feel strongly about the continued problematic issue of automated decision-making (algorithms) in the standardisation of student grades in 2020 - even as the GCSE process draws to a close.

<u>Topline.</u>

- Over 2000 State School GCSEs CAGs were externally moderated by an algorithm overseen by FFTeducationDataLabs in May 2020, prior to submission to the Exam Boards and OFQUAL in June 2020.
- This was called the FFT GCSE Statistical Subject Moderation - a free-service to ASPIRE subscribers allowing them to "validate" their proposed CAGS against the standardisation metrics defined by OFQUAL.
- Original CAG's were uploaded by the school, analysed by the algorithm and a report produced which benchmarked those grades and made recommendations to ensure compliance with OFQUALs guidance at the time.
- When OFQUAL discarded national standardisation on 17th August, grades reverted back to submitted CAGs, but at over 2000 FFT-moderated schools, those CAGs had already been subject to punitive standardisation against the OFQUAL metrics, leaving many borderline pass candidates to fail, despite their original CAGs predicting a pass.
- OFQUAL, rather than correcting this mistake in fairness to those students affected, have retrospectively changed their advice to specifically forbid appeals based on standardisation not happening.
- This is unfair and unjust to affect students but also risks a moral hazard in rewarding risk and punishing compliance.
- 30% of all pupils from state-funded schools don't achieve at least a grade 4 in Maths and English Language in typical years. Cohort level standardisation - as carried out by the FFT moderation in 2020, would have benchmarked pass rates for previous years in its report back to the school in 2020. Many of these borderline students will have moved from a teacher-assessed grade 4 to a standardised grade 3 fail as a result of the FFT intervention.
- Conformity Bias leads individual to do as they are recommended to do when they are directly responsible for the outcome

Background

In March 2020, the Minister for Education announced that GCSE and A levels would be cancelled, instructing regulator OFQUAL to develop replacement system. Their solution was based around Centre Assessed Grades (CAGs), a rank order of students in a subject cohort and - to ensure comparable outcomes with previous years - a standardisation of these grades at both national and centre level through the use of an algorithm. This would apply various metrics including Prior Attainment, Centre historic data and national historic data to "standardise results, both across this years' national cohort and relative to previous historic cohorts. As said, this was to ensure Comparable Outcomes for 2020.

OFQUAL provided very little initial guidance on how to achieve this, giving just broad instructions on how to formulate Centre Assessed grades based on historical data and prior performance. They

spoke in general terms of "ensure(ing) 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."

By August, after an outcry over systematic downgrading of CAGs by the OFQUAl/EAs, the algorithm was scrapped and original submitted CAGs were re-instated. In scrapping the algorithm, therefore, OFQUAL had also effectively abandoned Comparable Outcomes and by allowing CAGs to be re-instated, permitted grade inflation across all subjects and levels.

The Problem

In effecting a volteface over standardisation, abandoning the algorithm and reverting to submitted CAGs, OFQUAL inadvertently created a new unforeseen problem, entirely of its own making, which remains to the present moment and directly disadvantages an estimated 50% of all schools. By stating early on the process the metrics by which CAGs would be standardised and inviting Heads of Centre to include these when formulating and submitting their CAGs, OFQUAL essentially hard-baked this standardisation into the assessment process, with many centres (>50%) implementing the prescribed moderation prior to submitting their grades. This is hardly surprising given OFQUAL required Heads of Centre (HofC) to sign a statement which included the following condition: "I understand that exam boards will conduct a statistical standardisation exercise, using a model developed with OFQUAL, and that, if the profile of grades submitted is substantially different from what might be expected based on my centre's historical results and the prior attainment of this year's students, the grades for my centre will be adjusted to bring them into line with national standards." Given a natural Conformity Bias in relation to the regulator and Exam Bodies, many schools carried out pre-standardisation of their CAGS. When the standardisation Algorithm was discarded in August 2020, however, this pre-standardisation placed candidates at centres where it occurred at an instant disadvantage compared to centres which had only partially applied the moderation, or had ignored it entirely, prior to submitting their grades.

It is important to note that the standardisation metrics proposed - Prior Attainment and previous historical cohort data - did not apply to certain centre settings. Over 1000 independent schools entered no students for SATS in 2015 so have no standard of Prior Attainment to measure their GCSE cohort by. And many Free School will - by virtue of their recency - have no comparable historical centre data. So these two groups would obviously not have been able to pre-standardise their CAGs to the specification set out in OFQUAL's broad guidelines, not would their results - especially in small subject cohorts - be subject to standardisation at a national level by OFQUAL and the EAs. This group of schools would have submitted their original CAGs which, when the OFQUAL standardisation Algorithm was discarded, were upheld as their final grades. These are entirely unmoderated Centre Assessed Grades.

There is another group of centres, however, which DID possess the correct data to-hand to prepare their CAGs for standardisation and therefore satisfy OFQUAL's conditions at that time. Many state schools, having subscriptions to the FFT ASPIRE, a reporting and data tool for schools, local authorities and academy chains, were able to access the KS2 prior attainment data and historical cohort data for their centres from this source and potentially many pro-actively did. However over 2000 schools may have - in good faith - responded to an invitation from FFTEducationDataLabs to participate in an Algorithmic Standardisation exercise which, using the same metrics and data as the OFQUAL national Algorithm, pre-empted that standardisation and "hard-baked" into the final submitted CAGs. These CAGs should be considered "moderated" as they have had the algorithm applied to them, prior to submission.

The FFT process.
FFT developed the FFT GCSE Statistical Subject Moderation in late April and early May. In their invitation to schools, also disseminated via LEAs, FFT wrote "*FFT will provide a new service for*

secondary schools in May to help you validate and benchmark KS4 teacher assessment predictions. All secondary schools will be able to submit their teacher predictions to FFT in early May (whole grades only with no ranks). FFT will then provide a new 'Statisical (sic) subject moderation' report for all schools which analyses your predictions, factoring in FFT estimates and school subject performance in previous years. It will highlight any unusual subject level patterns which might merit further attention."

Any HofC, looking for assistance in formulating their CAGs - but also ensuring those CAGs conformed to the Government and OFQUAL's comparable outcomes prescriptive - was invited to "validate" and "benchmark" their CAGs, using a system they knew and trusted. "Validate" means 'prove the accuracy of' so many HofC would have anticipated the report they received from FFT would indeed help them meet OFQUAL's objectives.

Participation in the FFT Moderation service was free but required subscription to the FFT ASPIRE service, activation of Data Exchange and the submission of a .csv file containing the proposed CAGs, signed as accurate by the HoC

The FFT Algorithm was applied to this .csv data and a series of analyses and predictions were made from it by the Algorithm. These were provided in a report to participating centres but also shared on FFT Aspire and viewable by other centres. Many schools used this report to fulfil the stage 7 ASCL recommendations in their "Step by step approach to allocating grades using ASCL approach", "*Stage 7. Moderation of grades by senior leaders using national distributions and understanding of past centre performance.*" Some schools mention this FFT standardisation explicitly in their communication with parents, referring to "*externally generated data analysis from FFT Aspire*" (Cambourne School) and "*External Moderation (using FFT) In light of this the college will moderate TAGs before final submission using FFT*" (Henry Cort School). The sense is of centres "complying" with OFQUAL's requirements and guidelines, by utilising external data analysis via the FFT Algorithm. But as said before, the effect of this "external moderation" was to pre-standardise results, creating the "Moderated CAGs" mentioned above

The Outstanding Issue (at time of writing)

In total over 2000 schools participated in the FFT GCSE Statistical Subject Moderation in May 2020 (50% of the state school national cohort) and the reports FTT generated carried detailed suggestions on how original grades could be amended to "fit" both national and centre historic norms. This is EXACTLY the process the OFQUAL/Exam body Algorithm applied to all CAGs in June – moderated grades that were subsequently withdrawn or superseded by original CAGs. And yet their CAGs still carry that externally- moderated downgrade as it was "hard-baked" into the process in good faith. Good faith that has now created disadvantage for their cohorts.

In giving candidates the choice of their CAG or the OFQUAL moderated grade (whichever is higher) on 20th August OFQUAL is defining the same thing for FFT-affected candidates as their CAGs had already been moderated to "fit" with OFQUAL standardisation before submission. The original unmoderated CAG, was "lost"once submitted to external moderation by the FFT Algorithm in May. Candidates NOT affected by FFT moderation at other centres would have reverted to an original unmoderated CAG which had not been subject to moderation before submission. This is particularly pertinent to borderline pass candidates in GCSE Maths and English for whom a grade 3 downgrade from a Teacher Assessed Grade 4 would mean re-sits and potential impact on further education where any courses require a Grade 4 to progess.

OFQUAL, however, having created this issue itself by withdrawing the very national Algorithm standardisation that would have legitimated pre-standardisation of CAGs on August 17th, has instead penalised Centres that took this approach -consistent with OFQUAL guidance at the time by expressly refusing them leave to appeal - and reinstate their unmoderated, un-Standardised-by -the-FFT-Algorithm CAGs by stating: "*A school or college cannot raise concerns about its CAGs on the basis that another institution took a different approach, that different teachers could have come*

to a different judgement, or because the national process of standardisation did not operate as expected. Instead, the school or college would need to provide evidence of the original approach that it took and show why this was not appropriate, given the published guidance."

The solution.

Those 2000 centres that participated – in good faith – in the FFT Statistical Subject Moderation, will want to know why the original CAGs they uploaded in a csv file to FFT Aspire can't be reinstated given that the other 50% of schools that DIDN'T submit their grades have ended up having their original unmoderated CAGs accepted as final grades. The "risk" in not complying with OFQUAL guidance has been rewarded with higher grades and compliance has been punished. Therein lies a **moral hazard** that OFQUAL's current response wilfully ignores. And as usual, it is the "forgotten third" that suffered most as potential pass candidates in submitted CAGs were consigned by the automated FFT process to fail grades in order to fit a comparable outcomes metric OFQUAL itself discarded in August.

By recommending OFQUAL strike out the unfair and restrictive passage quoted above in their appeals process and allowing schools affected by the FFT Algorithm external moderation to re-instate their original CAGs as submitted to the service in May, the select committee will go some way to ensuring students at hose schools have a similar CAG to schools that, for whatever reason, didn't participate in the FFT Moderation before submitting their grades.

It is is a simple and sensible way to correct an obvious injustice and, having the csv file signed as authentic by the HofC before submission to the FFT service means no further validation would be necessary.

As OFQUAL have promised in their 17th August communication to schools and students "We have therefore decided that students be awarded their centre assessment for this summer – that is, the grade their school or college estimated was the grade they would most likely have achieved in their exam – or the moderated grade, whichever is higher." Removing external pre-submission moderation at affected schools is the final and necessary step towards this.

Conclusion

When OFQUAL state "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" by leading centres down a standardised comparable outcome routes, then abandoning that route and refusing to allow appeal by those centre that in good faith applied external moderation via the FFT GCSE Statistical Moderation Service, they have been manifestly UNFAIR to candidates who through no fault of their own were unable to sit exams, had Centre Assessed Grades carefully considered by those teachers that know them best. But then had those grades externally moderated by a version of an automated algorithm discredited by subsequent events.

In the light of this the Select Committee should consider the possibility of instructing OFQUAL and the Exam Authorities to allow appeal by schools on behalf of students whose original CAG, submitted by csv file to the FFT Moderation Service and, as a consequence, for those original CAGs to be re-instated.

Thank You

Mark Fretten.

Attached. Sample FFT GCSE Statistical Subject Moderation report, centre & subject levels.

FFT Statistical Moderation example report

Individual subject reports with 4 step review process:

1. Compare your KS2 prior attainment

2. Compare to last year's results

3. Compare to your own progress over the last 3 years (weighted)

4. Compare against FFT50 estimates (how *MIGHT* VA look)

September 2020