

For release on Monday 17th August 2020

## Ofqual grades algorithm: A recipe for unfairness

---

**Every year, on A-Level results day, there are thousands of disappointed students whose grades were lower than they hoped. The same was true in 2020, but rather than the results be based on grades achieved in examinations, they were based on a complex algorithm that gave different grades to students of the same ability and teacher-assessed grade, based on the size of their class/cohort, how their school peers had done in that subject in previous years, and the GCSE results the subject cohort had achieved.**

Many students from all backgrounds were awarded grades that are no reflection of their abilities and adversely impacts their future, but those from disadvantaged backgrounds were most affected.

upReach used the technical report Ofqual released, and historical grades databases for every school and subject for three years to investigate why. Capping student's performance based on how well their peers had done in the last three years was one issue - the impact of an unusually strong cohort another. And if the prior attainment (GCSE grades) of a student cohort was lower than usual - that probably got them all marked down too.

Driving the widening of the attainment gap was the fact that those in small subject cohorts were given their teacher assessed grade with no risk of downgrade. This was more likely in private schools (where there are more 'niche' subjects offered in small classes) where top grades went up 4.7%, compared to 0.3% at Sixth Form colleges who typically had bigger classes.

### **Commenting on the analysis, John Craven, Chief Executive of upReach and author of the research commented:**

*"Ofqual's flawed methodology resulted in rampant grade inflation in "private school" subjects such as Classics. By their own definition, Ofqual and the government have thus failed in their attempt at maintaining a "Gold Standard" by capping grades. Popular subjects more commonly studied at Sixth Form colleges saw no grade inflation. This, together with the use of a flawed algorithm based on historical data, has unfairly destroyed the dreams of thousands of ordinary students."*

*"The government need to urgently do the right thing, giving all students their teacher-awarded grades, rather than favouring private school students."*

## Methodology:

- The Ofqual Technical Report, released on 13th August, was thoroughly reviewed and used extensively. It provided a detailed description of the algorithm, together with selected data on outcomes for 2020 A-Level results.
- Preliminary analysis of the Report identified two trends that prompted investigation:
  - **A widening of the Attainment Gap** between those eligible and not eligible for Free School Meals, from 6.1% to 7.1% at A\*/A grade, reversing progress made to narrow this in recent years.
  - **Outperformance of Independent schools**, where A\*/A grades increased by 4.7%, particularly compared to **Sixth Form colleges**, where they rose only by 0.3%.
- Insufficient data was published by Ofqual to adequately explain the reasons behind these two trends, so other sources were identified to complement Ofqual data.
- Historical grade databases, previously published by the Department for Education, including grade data on every subject at every school in the country for the last three years, was used to complement the data provided by Ofqual.

## Summary of Key Findings:

There are three known major issues with the algorithm:

1. Firstly, the algorithm's reliance on historical schools data, meant that outstanding students at schools where few if any people had ever got top grades were unfairly capped, and have missed out on top university places as a direct result. This will have been one factor behind the widening of the Attainment Gap between students eligible for free school meals and the rest for A\* and A grades, reversing years of hard won progress by schools.
2. Secondly, this reliance on historical grades and on prior attainment data, meant the algorithm particularly disadvantaged students at schools that have been improving in recent years, had a particularly strong cohort in a subject this year, and/or that had weaker prior attainment data.
3. Thirdly, the algorithm had a bias that favoured those in small subject cohorts. This favoured smaller schools, and those doing niche subjects. Overall, this favoured private schools, where A\* and A grades rose 4.7%, compared to only 0.3% at Sixth Form colleges. The model created by upReach estimated that this "small subject" effect alone meant that Sixth Form colleges students were 20% more likely to have their teacher-assessed grade downgraded than private schools.

### **What drove outperformance in Independent schools compared to Sixth Form colleges?**

Based on the “small subject cohort” factor alone, our model estimates that A-Levels **grades at Sixth Form colleges were 20% more likely to be downgraded than those at Independent schools**. Other factors that require data not yet published by the DfE or Ofqual, are likely to make this an under-estimate. Our analysis suggests it is likely to be a combination of two factors:

- 1) The smaller independent schools having “small subject cohorts” for even averagely popular subjects such as Economics, meaning that they were more likely to have the teacher awarded grades used in the calculation, and less likely to be downgraded.
  
- 2) Rampant grade inflation in subjects more commonly studied by small cohorts at Independent schools and higher attaining state schools, such as Latin and Classics saw significant grade inflation, that was not found in other more popular subjects.

There was more easily accessible on the second factor, so this was our initial focus. In Classical subjects there were 7.7% more A\* and 10.4% more A\* or A grades than the prior year. This is most likely because in these subjects, students at almost all schools, were given grades that factored in teacher-assessed grades that Ofqual admitted were “over-optimistic.” In Latin, 97.4% of schools factored in teacher-assessed grades, in Classical Greek it was 98.9%, and in History of Art it was 81.2%. Analysis of our historical grades database showed that over 70% of schools offering Latin and History of Art A-Levels in 2019 were independent schools.

In contrast, subjects more commonly taught at Sixth Form colleges, such as Psychology, Sociology and Business Studies, saw little or no grade inflation (0.5% more students got A\* or A grades in Psychology, and just 0.2% in Sociology). Our findings suggest this is almost certainly because most schools (and almost all Sixth Form colleges) had their teacher-assessed grades completely ignored by the algorithm. In Psychology only 28.8% of schools factored in teacher awarded grades rather than just rankings, with just 35.2% in Sociology.

Around 30 times more students studied Sociology A-Level at Sixth Form and FE colleges than at private schools. Students studying these subjects at larger schools and colleges didn't benefit from the grade inflation evident in subjects more commonly taught at private schools or the most academic state schools.

On this issue, Ofqual stated in their Technical Report that “this is indicative of the modelling approach operating as intended.”

## How might GCSEs be impacted?

Ofqual plan to release a further detailed Technical Report that will enable a thorough analysis on results day, so this analysis is slightly limited in scope and required more assumptions.

Two different pieces of evidence both indicated that about 95-97% of grades awarded will use the algorithm, completely ignoring teacher awarded grades, relying only on their rankings. This compares to around 80%-82% of A Level grades. Based on the analysis, we anticipated the following issues:

1. Firstly, in normal times, ordinary non-selective schools can often see significant variation in the grades their students achieve each year, mainly due to how strong a cohort is. The reliance on historical data will mean that in schools with a strong cohort, particularly where their students have made more progress than usual, we'll see thousands of students downgraded from critical grade boundaries in Maths and English - e.g. by their teachers grades marked down from a 5 to a 4, or 4 to a 3. This will really penalise students at schools that have improved over the last few years, particularly where this cohort of students have made great progress.
2. Secondly, the bias still exists for "small subject cohorts" but GCSEs tend to be taught in bigger classes, so there are fewer students affected. Our analysis suggests that whereas at least three quarters of state school students will have all their grades calculated entirely by the algorithm, a typical private school student will have one grade where their teacher predicted grade is at least factored in. This will probably drive top grades rising at private schools faster than at state schools - just like they did at A Levels, just not as quickly.
3. Outside of English Language, English Literature and Maths, Ofqual's own Technical Report suggests GCSE grades will be less accurate than A-Level grades - with 10 out of 19 other GCSE subjects being less than 90% accurate (within +/- 1 grade) when modelling historical data.

**Limitations:**

- In much of the research, the lack of data published by Ofqual and limitations on the historical data available meant that reasonable simplifying assumptions were required, particularly in the analysis on A-Level data. For example, where a school has 5 or fewer entries for an A-Level subject, the DfE historical grade databases suppressed both the total entries and the grades achieved (labelling them “Supp”).
- The statistical methods used by Ofqual, involving concepts such as the “harmonic mean,” created challenges in clearly communicating key definitions such as a small subject cohort to stakeholders, and there was often incorrect analysis, particularly on social media as a result. Teacher assessed grades were awarded without adjustment where the harmonic mean was 5 students or less, but this was wrongly interpreted to mean 5 or fewer students doing the exam in 2020. In the absence of data from Ofqual, our analysis necessarily made assumptions about how the harmonic mean.
- The Ofqual Technical Report was 319 pages long, highly technical, and only published four days before the date this research is published (the urgency of the situation for many students necessitated publishing this analysis at the earliest opportunity). We do not anticipate any meaningful errors or omissions were made but there has been insufficient time for the level of validation that would normally be preferred.
- Best efforts were made to ensure all assumptions made were reasonable.
- We call on Ofqual to publish more detailed data that can be an authoritative source for future analysis and evaluation of the algorithm.

**Author:** John Craven, Chief Executive, upReach

## **Information about upReach**

**upReach** are an award-winning social mobility charity that support over 1,600 students to secure top graduate jobs, in partnership with leading universities and employers such as Bank of America, Civil Service Fast Stream, Deloitte, Goldman Sachs, McKinsey and Slaughter and May. A team of 34 employees provide personalised support helping to broaden horizons, raise aspirations and enable disadvantaged students to develop the skills, networks and experiences to succeed on merit. upReach won Charity of the Year in 2019 at the Charity Times Awards, and won the 2020 Embracing Digital Award in May's Charity Governance Awards. Learn more at [www.upreach.org.uk](http://www.upreach.org.uk)

## **Contact details:**

John Craven, Chief Executive, upReach      [john.craven@upreach.org.uk](mailto:john.craven@upreach.org.uk), 07971 274469

## **CovidCohort Grades Evaluation platform:**

At midnight on 12th August, amidst fears that high attaining disadvantaged students would be most at risk of having their grades downgraded, upReach launched their **Grades Evaluation platform** for students. Full details are [available here](#), and the platform can be accessed [here](#).

## **The CovidCohort Grades Evaluation platform was designed to enable students to:**

- Better understand the likelihood of whether their grades might have been capped due to Ofqual's standardisation process.
- Find out how their A-Level grades compare to those awarded in each subject at their school/college in prior years.
- Discover if they're in a group where grades are historically under-predicted.
- Learn how their socio-economic background may have impacted upon their ability to achieve the same high A-Level grades as others with the same academic potential at their preferred university.

## **The free four-page personalised Grades Evaluation included:**

1. **A-Level Grades Evaluation:** At their school, in the prior 3 years, what percentage of students got each grade for the subject they studied - e.g. 25% got a grade C, and a comparison to national averages. Their grade will be shown, allowing a "within-school" comparison to be made, analogous to a ranking - e.g. "top 15%." This is the most critical information used by Ofqual to determine grades, alongside teacher rankings.
2. **Flag for Likelihood of Lowered Grade:** Indicating the likelihood that their grade was capped / lowered (which would be red, orange, yellow or green, depending on factors including how many students previously got top grades at their school)
3. **"Statistical Outperformer" Flag for risk of Under-predicted grade:** Indicating whether the student was in a demographic group that statistically outperforms their teacher predictions (based on the latest academic research).
4. **Contextualisation Adjustment Score:** Based on 14 different indicators of socio-economic disadvantage, this score indicates the extent to which a student from their background may be unable to realise their academic potential in a "normal year," for example, due to attending a very low performing school. upReach's contextualisation platform, REALrating, has used this to generate contextualised A-Level grades since being built in 2017, and is used by employers to identify hidden academic potential. Most students have a score of zero, but for others this gives a boost of up to 3 grades.

## Appendix - upReach's REALrating and Contextualised Adjustment Score

**REAL (Relative Educational Attainment Level)** is a contextualisation platform built by upReach in 2017 to help employers identify hidden potential. A research and evidence-backed algorithm takes into account 14 different indicators of disadvantage to calculate a net indicator of socio-economic disadvantage called a Contextualised Adjustment Score. This is used to calculate contextualised A-Level grades, which are used by employers as a better indicator of academic potential than A-Level grades alone. Further details, including student case studies, can be found [here](#). Learn more at <https://realrating.co.uk/>

Many universities acknowledge that A-Level grades are a function of both academic potential, and a student's opportunity to fulfil their potential. They recognise that some people from disadvantaged backgrounds do not have the same opportunities to fulfil their potential at school, and that their A-Level grades may not fully recognise their potential.

Research also shows that various socio-economic factors impact the likelihood of students with the same A-Level grades getting a good degree. For example, on average, those from state schools are more likely to get a 2.1 or first class degree than those from private schools with the same grades.

As a result, some universities reduce their standard offer, for example, from AAA to AAB or ABB, for students that meet certain criteria.

Since 2017, upReach have calculated a "Contextualisation Adjustment" score based on 14 different indicators of socio-economic disadvantage, that indicates the extent to which a student from their background may be unable to realise their academic potential, for example, due to attending a very low performing school. upReach's contextualisation platform, REALrating, has used this to generate contextualised A-Level grades since being built in 2017, and is used by employers to identify hidden academic potential. Most students have a score of zero, but for others this gives a boost of up to 3 A-Level grades.

### Utilisation:

In years prior to 2020, the REALrating gives students a better indication of their academic capabilities than A-Levels alone, and enables them to see how their academic potential compares to their peers, since it shows performance in context of their background and schooling.

Universities give lower "contextual" offers to some students to recognise this unrealised academic potential. Graduate employers are increasingly looking beyond A-Level grades to assess academic potential. REALrating provides Employers a fairer measure of applicants' academic achievements than A-Levels alone.

### Methodology:

REALratings are produced using a sophisticated model that takes into account factors such as school type and background. The methodology is based on a comprehensive review of the published research into how background and school type influence academic performance and analysis of how universities and employers are using contextualised data to inform admissions and recruitment.

The REALrating is calculated by adjusting the number of UCAS points attained according to a number of different socio-economic factors - this number becomes the "REALrating."