



## Exam Results, Fairness, & Trust: A response to

[Karen Lancaster](#)

Mary Richardson

\*\*\*

*This is the second post in a PESGB Blog series focussed on educational assessment and the UK Government's handling of England's 2020 national exams (GCSEs and A Levels) in the wake of Covid-19.*

\*\*\*

The grades calculated in summer 2020 were indeed problematic, but it is not the algorithm that is central to this problem; it is politics. In a normal year, the use of [comparable outcomes](#) to guide the awarding of exams means that we might expect a similar cumulative percentage of students to sit between each of the grade boundaries and this is why the actual score for the boundary changes year on year. This is the standard and the awarding bodies (exam boards) use an algorithm to model the data based on a range of variables that impact student performance including sitting the exams themselves.

The purpose of these algorithms is to *moderate* the marking process and enable us to make a reasonable national judgment of outcomes. It's not



perfect; it's riddled with error! But it is one of the best ways to manage large data sets that are a proxy for student achievement.

Every year, all schools submit data (with the *exception* of teacher estimated grades), to the respective awarding bodies and these data are fed into a statistical model that is viewed alongside the actual examination results so that decisions about where to place the grade boundaries can be made. Many of us knew the application of the usual judgement algorithm would lead to a skewing of outcomes, simply because the awarding bodies and Ofqual were not able to look at data sets similar to those of preceding years. Experts spent many months creating a variety of models in attempts to simulate data as best they could – but this was an impossible task as they were creating something very new in a short space of time. It was a political decision from central government to use this method – there was no pedagogical knowledge or consideration of fairness in that decision. [Mr. Williamson et al were firm in their wish to 'maintain the standard'](#) (see also [here](#)). Of course, what the initial grading revealed was that it would not be possible to maintain the standards because the information needed to do that was incomplete. On Results Day it emerged that almost 40% of students received very low grades and of these, it was [students from the lowest socio-economic backgrounds](#) who were hit hardest – a triple COVID whammy of disrupted schooling, lack of home resources, and, to top it off, poor exam results.

Assessment research shows us that in state schools, teachers tend to [over-predict](#) student grades (their counterparts in independent schools don't do this, they are much more accurate in their predictions). The reasons for over-prediction are complex – it is, we think, to do with accountability measures in state schools, pressure on teachers from Senior Leaders, attempts to motivate students coupled with the impact of terminal test outcomes on school rankings... Such factors combine to make results eye-wateringly high stakes.

Going forward [it is confidence in educational assessment that matters](#) – there are several things we can learn from these experiences. Firstly, perhaps the cohorts of 2020 and 2021 will *not* fare so well given that they are already being labelled the COVID generation and their results are being viewed with suspicion. Secondly, we should ask: [why do we cling to such an archaic national assessment system](#)? We would not be facing this continual problem if we had not decided to go with an all-or-nothing approach (exams) to characterising schooling. And finally, we need to think very hard about why teachers over-predict.

I disagree with [Karen Lancaster](#) that the trend of over-prediction doesn't matter. I think it does – it is actually the root of the perceived unfairness in national testing, so why do we continue to accept it? If England had retained at least some modular components for A levels and GCSEs, then there would at least have been more secure evidence – including coursework – on which to



base students' achievements. These pressures are not just damaging for teachers, they continue to embed inequity for students.