

## Grading and Ranking – a field guide

While the final details regarding the post calculation processes that awarding bodies will be doing to moderate student grades are yet to be finalised, we do now have a clear understanding of what schools and colleges will be required to do as part of the ‘calculated grades’ methodology in this year’s awarding process. For us to navigate this process effectively, and give our students the best possible chance of getting the grades they deserve, we need to go into it with a clear understanding of what makes this process particularly difficult and a sense of the implications of the decisions we make. We need to take the opportunity to complete as much well-judged quality assurance as we can, as a fundamental part of the process. The potential for getting it wrong is significant. Just remember that some schools have already emailed students with their final grades – quite a feat prior to the issuing of any agreed process regarding their awarding.

### **The challenge of awarding in the absence of an exam season**

On the face of it, what schools and college are required to do is a straightforward exercise: subject teams need to think about each student in turn, considering the work they have produced over the course, the quality of their work in class, the trajectory of improvement (or otherwise) they have been on; they then need to place this knowledge of the individual student against what they know of all the students they have ever known attempting that particular qualification, drawing on their experience of grading over the years and (ultimately) decide on the grade that they think the student should be awarded.

While the awarding of grades in a centre is relatively simple, consider how different this is from the awarding of other centre assessed marks such as those for coursework. What makes awarding so complicated in this instance is the absence of a standardised and moderated process by which marks (and in turn grades) are awarded across all centres in normal times. Awarding bodies have processes that are remarkably effective (though of course human and imperfect) in ensuring the right students get the right outcomes, but this year awarding has to take place in the absence of these systems. Awarding bodies will have to develop some system of moderation that ensures that the grades awarded this year are in line with those from previous years. We know from UCAS predicted grades that there is a degree of generosity in college predictions, and there is good reason to suspect that there will be generosity in calculated grades also, so there will need to be a moderation process. One of Ofsted’s observations in moving away from looking at centre produced evidence of student progress was how optimistic in-year judgements tend to be. He was probably not exaggerating when one HMI commented that in thirteen years of inspecting, he had never encountered a centre which had concluded that it was going to be a worse year for results than the year before.

We need to establish why moderation is challenging in this context, because if we understand the problem we have a better chance of producing a meaningful solution. Just why is it so difficult to

settle upon an appropriate grade profile for an individual centre? What is really boils down to is the absence of a consistent evidence base which can identify how students are performing in different centres. Consider the following:

- Some centres have better qualified candidates than other centres. One thing decades of value added analysis has shown us is that there is a relationship between prior attainment and student outcomes. It is important that prior attainment is taken into account when deciding on the appropriate grades for a centre, but it is by far from the whole story.
- While there is a real advantage of recognising prior attainment, there is a danger of sticking to tightly to what a value added model would predict for a centre. Some centres are better than others, and in each centre some departments are better than others. To what extent will historical performance be taken into account in moderating the performance of individual centres? There are subject departments in sixth form colleges which have outperformed predictions based on prior attainment by at least half a grade per student for each of the last ten years. If their outcomes were limited to what a regression line would predict, around half of their students would get a grade below what they would have achieved had the exams gone ahead as planned. These students would clearly be disadvantaged (contradicting the stated aim of the whole process).
- So we need prior attainment to be taken into account, and we need a sense of historical performance, but this may still provide a flawed basis for analysis. Some centres are on a different trajectory to other centres, and in each centre some subjects will be on different trajectories. Imagine a department which had a particularly poor year in 2019. If it had made no real effort to make any changes, then historical performance may be a useful starting point for grading. But what about one where the head of the underperforming department had been replaced, and the new leader had formed a new dynamic and highly effective team. It would be unfair if that department was pegged to performance the previous year. But there is a danger in considering trajectory in that virtually no-one believes themselves to be on anything but an upward trajectory. Sensitivity to genuine trajectory will be hugely challenging for the awarding bodies.
- If prior attainment and prior performance are to be taken into account, how can they be applied? Given the multiplicity of awarding bodies, it only really makes sense to conduct the analysis at individual subject level in each subject. It is entirely sensible to use prior attainment and prior performance as part of the process, but getting this right is a real challenge, and it is not an area of expertise that is particularly well developed at the awarding bodies. While awarding bodies are skilled at navigating comparable outcomes at qualification level, this is not something that they routinely use at subject level with individual school and college cohorts.
- Not all centres will have any 'history' to draw on in establishing prior 'raw' performance, or performance related to prior attainment. Some recently opened centres will be in their second year of operation and will have no history of performance to go on. More commonly, centres will be new to the particular board they are with that year, so will have no history to consider in moderation.
- We should not forget that we are only just coming out of A level reform, and for a significant number of subjects (notably mathematics) we only have one year of 'history' to go on. We know from the 'sawtooth effect' that it is likely that in the first few years of a specification, the quality of students' work improves, and there is a tendency for the marks required to achieve particular grades to rise. We don't know the extent to which everyone else is improving.

Ofqual's approach promises a moderation process that takes in to account prior attainment and prior performance. There is a real potential in the system for grade inflation here as centres try to push the best possible grades for their students. There is in turn a very real danger that the processes used by awarding bodies to dampen the inflation and restore the national grade profile to a credible one may dampen performance in centres that were entirely legitimate in their awarding. Colleges will need to be very careful in their calculation of grades.

## Calculating grades

In the calculated grades process, there are two things that schools and colleges are being asked for: grades for each subject and a rank order for each subject. Both of these assessment decisions are significant. In short, the grade will determine what grade students might get, and the rank order will determine what grade they do get.

For big centres, and therefore for almost all sixth form colleges, ranking is an issue. A sole teacher can at least apply a consistent methodology to deriving rank order (however imperfect the process). As soon as more than one teacher is involved, the question of standardisation arises. The exact position a student holds in a cohort ranking has a consequence. While a college may award a grade A for a particular student, the moderation process may judge that a college has awarded too many A grades, and will use the ranking to identify which students are to drop down to a B. There will be cases then, where the exact position in a college ranking will determine whether you get an A or a B grade. For a few students, these will be life changing or life making decisions. We need to ensure that all colleagues are aware of the significance of a student being towards the bottom of a ranking of a particular grade.

Colleges will need to define what information they will use in each individual subject to establish the calculated grade and rank order. I would suggest the following things will form part of the overall understanding:

- The quality of student work (in all its forms). Those departments which are relatively small, or those that have a common programme of assessment will be at a significant advantage here.
- Work completed since the closure for coronavirus should not play a significant role in grading or ranking decisions, as work produced at home is more sensitive to the context of its creation than would produced in a college environment
- Information about the subject specific target grade for each student
- An understanding of how students perform nationally in the subject concerned, and the relationship with prior attainment.
- How students performed in the centre in question in previous years, and how this year's cohort compares to historical cohorts in terms of prior attainment.
- We need also to sense check the cumulative impact of our grading decisions in terms of overall performance, and in terms of particular equality and diversity characteristics (particularly gender).

Target grade setting systems provide useful information to help guide colleges through this process, but the guidance from Ofqual discourages the uncritical use of target grades. There are a couple of points that are vital to consider when using target grades as a sense check:

- Note that some target setting systems give the same target regardless of the subjects a student is taking. This is done for good reasons. For example, a college may decide that it

does not want to have to explain to parents why a student has three different target grades. When it comes to conversations with students, having one target makes sense, but when it comes to understanding the grade a student is likely to achieve in a particular subject, a single target grade can be very misleading. We know from the national data that students perform much more highly in some subjects than peers with similar levels of prior attainment attempting other subjects. It is no secret that likely outcomes in Film Studies and Photography are much higher than they are for Physics and Biology. Using a single target grade runs the risk of misleading colleagues as to the grades that students are likely to achieve.

- We need also to spend a little time looking at the notion of 'likely'. Target grades are not an exact science, and at every level of prior attainment a spread of grades is likely, and the target grades represents (usually) the mid-point of likeliness. A student may have a target grade of a C, but a significant number of students with that target will get a B or D, and some will get an A or an E. Nationally, around 60% of students equal their target grades, 20% beat them, and 20% fail to meet them.
- The value of any target setting system is to provide a starting point for discussion about a student. Their actual performance may be some way away from the target. The target grade provides a useful sense check, but we should not attribute a spurious accuracy to it. Target setting systems can also give us a (very) provisional rank order to work on.
- Colleges will need to consider where they are in relation to national average performance when considering how to interpret target grades. Put simply, if a college has consistently out-performed similarly qualified students nationally (as evident in value added scores) it should not limit calculated grades to what students achieve nationally.

To help colleges, the six dimensions analysis has developed a 'mathematical grades' target setting report. The targets work at individual subject level, and are calculated at national average performance. All colleges can receive a free of charge mathematical grades report, to help them consider the appropriate grade for each student in a subject. The report is a spreadsheet output with a row for every A level enrolment, giving the subject target grade for the student, adjusting target grade in light of the prior attainment of the student and patterns of performance in the subject concerned. The benchmarks are calculated using the 2019 data, which is particularly important for subjects such as mathematics and law, which saw significant changes in national performance in their first reformed outcomes. To receive a mathematical grades report, colleges will need to complete a data return. Data collection instructions will be issued in the next few days and the reports will be returned to colleges before the end of the (as was) Easter break (17/04/20).

## **The task ahead**

- Colleges need to develop a timeline for proposing, checking and submitting calculated grades and rank orders to deadline (not earlier than 29/05/20).
- Colleges need to communicate with students and parents regarding the process, being absolutely clear that in order to maintain the integrity of the process, centres are expressly forbidden from sharing proposed grades or rank order positions with students or parents (and before they ask, GDPR does not extend any rights to them in respect of this data)

- Colleges need to consider what appropriate evidence will be in each subject. This will vary from subject to subject, according to the availability of partially completed and completed coursework, mock exams, common pieces of work, and so forth. Ofqual appear to be quite flexible about what this evidence base will be as long as it leads to a “fair, reasonable and carefully considered judgement” of the grade the student would be most likely to have achieved had the exams gone ahead as normal. Predicted grades from the UCAS process should be used with extreme caution, given they were set the best part of a year ago, and tend to offer a rather optimistic view of likely performance. However, they could provide a useful starting point for discussion.
- It is important that the evidence base is one that works for the individual subject (and does not have to be a college wide approach, and probably should not be), but it is vital that the basis for judgement is understood by all. Note that in coming to a fair consideration, access arrangements should be considered.
- Colleges will need to put in place processes to check the grades that have been provisionally awarded in individual subjects, check the grade profiles that emerge, and check that there are no systematic distortions in the data (by gender, for example). Bear in mind that two people will need to countersign the declaration confirming that the judgments meet Ofqual requirements. The Head of Centre will have to complete an overall declaration for the centre. Centres will not need to retain the evidence itself (in contrast to Cambridge International qualifications)

There are, of course, other challenges ahead. Not least of these is dealing with incoming students who have had grades awarded using this process, and may well not have had the standard induction process of welcome days, etc. Some will not wish to accept the grades they have been awarded, and may wish to take up the options of Autumn Term exams. Colleges will need to be prepared to address their disappointment and provide appropriate support and guidance.

In the short term though, getting this process right is a huge challenge. It requires deep thinking about students and learning, and it may (probably will) involve huge logistical challenges, particularly for large departments working at distance with limited access to student work. But it is a process that we can get right: by deriving appropriate rank orders, and assembling grade profiles which are credible given historic performance at the centre, and the prior attainment of the students we deal with.

Nick Allen, 31/03/20