

A Commentary on the use of Calculated Grades for the Leaving Certificate

Michael O'Leary

CARPE/DCU

May 13, 2020

Note: Apart from the inclusion of reference list below, this article is the same as the one that was published in the Irish Times <https://www.irishtimes.com/news/education/would-leaving-cert-students-estimates-of-their-own-grades-help-teachers-1.4252424>

The cancellation of Leaving Certificate (LC) exams was a difficult decision for all involved. Up until Friday May 8, Minister Joe McHugh and his officials at the Department of Education and Skills (DES) were between a rock and a hard place. They knew that planning to go ahead with the Leaving Cert exams over the summer was risky given the health threat; logistically challenging, given the requirement for social distancing in exam centres; and open to criticism, given the range of challenges being experienced by students with poor internet access and/or with high levels of anxiety. But they were also aware of the difficulties involved in conceiving of an implementable alternative to Leaving Certificate exams, with the legislative constraints of the State Examinations Commission being involved in any alternative arrangements being a major one. For the teacher unions, the cancellation of the exams meant crossing a red line by agreeing that teachers could be involved in a process of assessing their own students for certification purposes. For representatives of student, parent and management bodies it meant advocating for positions that were not universally supported by members.

So what now? Well why not begin by placing our trust in the professional judgement and expertise of teachers in the way we have done with our medical practitioners? It's true that Irish research evidence on the accuracy of teacher-predicted grades is lacking and that research conducted with teachers in the UK has been presented as evidence in arguments made against the practice. However, contextual issues in the research have been missed in some reporting and should be noted.

Predicted grades are used by the CAO equivalent in the UK, UCAS, to make *provisional* university place offers to students, a situation that is very different to where we in Ireland

now find ourselves. The 16% accuracy rate reported in a UK study (Wyness, 2016) and quoted in some media reports relates to predictions involving a combination of three A level subjects, not one. In addition, what hasn't been explained clearly enough is that about 75% of the predictions were *over*-predictions (when compared to the actual results) as a consequence of many teachers using the process to motivate their students in the period before exams taking place. Interestingly, over-prediction was as likely occur to with the majority of disadvantaged students as it was with their non-disadvantaged counterparts. However, the fact that very high-achieving students in disadvantaged schools were likely to be under-predicted (about 3,000 students over three years) should be noted.

The findings from individual studies conducted over the past ten years in the UK, New Zealand and from a meta-analysis of 75 studies from the United States and various European countries suggest that the correlation between teachers' judgments of students' academic achievement and students' actual test performance is positive and fairly high (e.g. Sudkamp et al., 2012). The correlation coefficient is around 0.6 - a finding that, in essence, indicates that there is a similarity between how teachers and tests rank-order students, but the rankings are not always the same. This arises from the reality that teacher judgements and standardised exams/tests are different assessments. Importantly, we must not assume that the exam/test ranking is the correct one - I will return to this issue. There is also evidence in the literature indicating that we need to pay particular attention to the accuracy of teacher judgements about students from low socio-economic backgrounds and/or with low achievement levels (e.g. Meissel et al., 2017; Murphy & Wyness, 2020). Overall, it is important to note that much of the research is based on individual teacher expectations about *future* performance and does not necessarily relate directly to the accuracy of judgements arrived at when teachers apply clear criteria and work with colleagues and their school leaders to arrive at a decision about *current* achievement. With that in mind, the change of terminology from *predicted grades* to *calculated grades* is useful in so far as it helps to remind us that the focus now should be less on trying to replicate a LC result and more on how informed teacher judgements can be used to arrive at the best possible measure of student achievement.

While the LC examination system has many strengths, we should be careful about holding it up as a paragon of accuracy and fairness. There are very good reasons why efforts to

reform it are underway. The reality is that there is no ultimate truth in a LC result (or the outcomes from any assessment for that matter) because each exam cannot measure all elements of a subject area. As a consequence, every educational assessment contains what is called *measurement error* (which also accounts for some of the variation in rankings). It is analogous to the idea of the *margin of error* (e.g. $\pm 3\%$) in opinion polls that involve a sample rather than all possible respondents. There are also the myriad factors that affect student performance on the day of an exam, e.g. misreading a question, not feeling well etc. The LC is fair in so far as everyone takes the same test and under the same conditions that include anonymous marking. The public has confidence in the system and that is important. However, students do not arrive at the testing centres with nothing but ability and a track record of diligent study separating them. Some had better teachers than others and some were able to avail of the benefits that economic advantage bestows e.g. grinds. Indeed, so many of the problems faced by students over the past while in terms of having access to technology, having a quiet space at home to study and so on, have been relevant to the LC fairness issue long before the arrival of Covid-19.

Needless to say, the guidance provided by the DES to teachers and schools will be crucial in helping them make the best possible decisions about student marks and class rankings as well as how to handle conflicts of interest. At the very least, the collaborative nature of the work should mitigate the danger of individual teacher biases coming into play. What to do in situations where prior information for individual students is inadequate and/or where students study a subject outside of school remains problematic. Both issues prompt me to wonder if the addition of estimated grades (or marks) submitted by all students for the subjects they were planning to take, along with a justification for each using a DES approved *pro-forma* would have helped the decision making process in schools. I acknowledge that the idea of students predicting their own grades is unusual and that research tells us that students are prone to overestimate their grades (e.g. Attwood et al., 2013). However, the idea is democratic and, as well as providing additional information for teachers, would have indicated that we trust our young adults to be responsibly involved in a process that will have a major impact on their futures. It might also have been a way of guarding against the possibility of canvassing by students and their parents/guardians.

In the guidance provided to teachers and principals it would be important to include the conclusions from the research on teacher judgements, especially those that speak to fairness for disadvantaged students. Two other points are also worth highlighting. Over the years a number of studies linking Junior and Leaving Certificate data have been conducted at the Educational Research Centre, (e.g. Millar and Kelly, 1999) and, assuming issues of data protection can be addressed, a new study undertaken in the very short term would provide robust data for teachers to use. We also need to remember that, over the past number of years, many teachers have worked with colleagues during *Subject Learning and Assessment Review* (SLAR) meetings to assess student work as part of Junior Cycle reform. This experience is likely to be very useful now as they go about the process of calculating their students' marks/grades.

So let's row in behind this plan and give it every chance to succeed. While every student deserves to be treated fairly, so does every teacher. Students have the right to appeal and, if unhappy with the outcome, can take an exam at a later stage. I'm also confident that further and higher education institutions will make every effort to enhance access schemes if required. Teachers, like their medical counterparts, have the right to make professional judgements without fear or favour, but I worry that headlines that appeared in British newspapers in April will begin appearing here (e.g. *Parents and pupils overwhelm schools with pleas for good grades* – The Guardian, 19/4/20).

All of this is not to suggest we should not leave our critical faculties behind. Far from it. I understand the need to review grades submitted by schools and to apply statistical procedures in some cases to ensure "common national standards." However, we should remember that there is no objective truth to be found in distributions of grades from previous LC years either and that conclusions drawn from aggregated data (as opposed to student level data) can be problematic (see Gilleece, 2014 for more detail). This will be especially important to bear in mind when schools are making strong cases for awarding individual grades/marks that may not fit an established pattern. Moreover, planning for a programme of research should begin immediately to evaluate the extent to which this pandemic has forced us down a path that may or may not prove useful in reforming what we do at the end of secondary education in Ireland.

References

- Attwood, G., Croll, P., Fuller, C., & Last, K. (2013). The accuracy of students' predictions of their GCSE grades. *Educational Studies* 39(4), 444-454.
DOI:10.1080/03055698.2013.776945
- Gilleece, L. (2014). Understanding achievement differences between schools in Ireland – can existing data-sets help? *Irish Educational Studies*, 33(1), 75-98.
- Meissel, K., Meyer, F., Yao, E.S., & Rubie-Davies, C.M. (2017). Subjectivity of teacher judgments: Exploring student characteristics that influence teacher judgments of student ability. *Teaching and Teacher Education*, 68, 48-60.
- Millar, D., and D. Kelly. 1999. *From Junior to Leaving Certificate: A longitudinal study of 1994 Junior Certificate candidates who took the Leaving Certificate Examination in 1997: Final report*. Dublin: ERC/NCCA.
- Murphy, R., & Wyness, G. (2020). *Minority Report: the impact of predicted grades on university admissions of disadvantaged groups*. London: University College London, Centre for Education Policy and Equalising Opportunities (CEPEO)
- Sudkamp, A., Kaiser, J., & Möeller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762.
- Wyness, G. (2016). *Predicted grades: accuracy and impact. A report for University and College Union*. London: University College London.
- Michael O'Leary holds the Prometric Chair in Assessment and Directs the Centre for Assessment Research, Policy and Professional Practice (CARPE) at the Institute of Education, Dublin City University.