ROYAL STATISTICAL SOCIETY

DATA | EVIDENCE | DECISIONS

RSS EPAG STATEMENT ON GRADE ADJUSTMENT IN 2020 EXAMINATIONS IN THE UK

6 August 2020

We have now seen the release of some data and information from the Scottish Qualifications Authority[1] about how they adjusted teacher-predicted grades; next week we expect to see similar information from Ofqual and devolved qualification authorities about exam results in England, Wales and Northern Ireland. In the light of these, the Royal Statistical Society[i] (RSS) and its Education Policy Advisory Group[ii] (EPAG) sets out some observations.

We stress that these are not views arrived at in hindsight – we have been raising these issues from the earliest Ofqual consultations[2], and in our evidence to the House of Commons Education Select Committee[3].

We start by noting that actual exam results themselves are an imperfect measure of what students know. There is variability by which exam questions are set and answered, there is marker variability, and individual student performance can vary day to day. So we should not hold the system in this year of estimated grades to a higher standard.

From the outset, the RSS has raised questions about the proposed system for statistically adjusting ('moderating') the teacher-estimated grades. We have grouped these under three main statistical issues and one principle that relates to the process of statistical adjustment overall.

First, it was always likely that teacher-estimated grades would be optimistic; we have long known this was the case in, for example, teacher-predicted grades for UCAS[4]. This systematic uplift (treated by statisticians as 'bias') could have been mitigated by requiring in addition submission of individual-level data about prior exam attainment, performance in coursework and mock exams. Having this individual-level data might have allowed consideration of individual adjustments, rather than taking an approach based mainly on student rankings and historic performance of exam centres.

Second, the RSS raised concerns about the use to be made of teacher *rankings* of individual students by subject within exam centres. As RSS Fellow, Professor Guy Nason has pointed out[5], these rankings are subject to some uncertainty, particularly so for middle-ranked students. Again, there are statistical techniques for taking account of uncertainty around rankings.

---

[1] Scottish Qualifications Authority Technical Report, National Qualifications 2020 Awarding – Methodology Report, August 2020, accessed 5 August 2020, https://www.sqa.org.uk/sqa/files_ccc/SQAAwardingMethodology2020Report.pdf

[2] Royal Statistical Society news story, RSS alerts Ofqual to stats issues relating to 2020 exam grading, 6 May 2020, accessed 5 August 2020, https://rss.org.uk/news-publication/news-publications/2020/general-news/rss-alerts-ofqual-to-stats-issues-relating-to-2020/

[3] Written evidence by the Royal Statistical Society to the Education Select Committee, The impact of COVID-19 on education and children's services inquiry, 8 June 2020, accessed 5 August 2020, https://committees.parliament.uk/writtenevidence/6407/html/

[4] Written evidence by Professor Guy Nason to the Education Select Committee, The impact of COVID-19 on education and children's services inquiry, published 22 July 2020, accessed 5 August 2020, https://committees.parliament.uk/writtenevidence/8594/html/

[5] Written evidence by Professor Guy Nason to the Education Select Committee, The impact of COVID-19 on education and children's services inquiry, published 22 July 2020, accessed 5 August 2020, https://committees.parliament.uk/writtenevidence/8594/html/

Third, the RSS raised questions about the *variability or volatility* of exam centre results year-by-year, as these can vary by intakes, type of exam centre, steps taken by centres, and so on, especially for smaller subjects or intakes. An exam centre with steady average or median grades might still show some variation in the overall distribution of grades within a subject, year on year.

These are complex statistical issues. The RSS was clear from the beginning that we understood why deferring exams at a time of such disrupted teaching had disadvantages and could also be unfair (since different schools were disrupted differently), and that many young people did not want to delay their future plans. However, the overarching principle by which we would have approached this would have been based on greater transparency about the statistical adjustment being considered. We have consistently raised issues about getting more accurate and detailed data on which to consider various ways of making the statistical adjustment, and the need to discuss options more openly before a final method was applied. Any statistical algorithm embeds a range of judgements and choices; it is not simply a technically obvious and neutral procedure. Calibrating this year's estimated grades to previous years' exam results is one such choice. How to take account of evidence of individual students' prior attainments is another. How to take account of uncertainty is another.

So from the outset we have urged Ofqual to be more transparent about the choices it would have to make, and the quality and certainty of the data on which these would have to be based. We understood concerns that the model could not be tested until the data were gathered (though we think a fuller range of data might have been sought), but in the months after the teacher-estimated grades and rankings were submitted, we called for a wider public discussion.

This process may not have resulted in consensus, but it would have resulted in a discussion about statistical adjustment *before* individual grades were issued. In addition, it might have given information (which is not, in Scotland at least, available even now, since the detailed statistical algorithm has not been released) about which categories of students might have been more affected by the algorithm and who might therefore have more powerful grounds for appeal.

Instead, only very limited information[6] has been released in advance about the statistical adjustment procedure, or why the decision was taken to treat exam centre historical performance as the anchor for adjustments, or whether individual student historical performance on previous exams or coursework or mocks gathered before Covid-19 disrupted teaching could have played a role. Arguably, the integrity of the system has been upheld, but without detailed information on which to judge 'fairness' to individual students.

As the RSS wrote to the House of Commons Education Select Committee, we do not believe that the current system can be proof against appeals. But in the absence of the statistical details, neither we nor anyone else can really be sure about whose grade estimates might have been most affected, or – more to the point – whose were most 'unfairly' adjusted.

[6] Professor Guy Nason's personal comments on the SQA Technical Report: National Qualifications 2020 Awarding - Methodology Report, August 2020, Publication Code BA8262, 4 August 2020, accessed 5 August 2020, http://wwwf.imperial.ac.uk/~gnason/Pers_SQA_Aug4.pdf

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

We write this with regret, as we know the pressures the exam authorities have been under, and we can understand why a statistical adjustment, or 'moderation', would be necessary. But we persist in our view that more transparency earlier on would have been better – for students and their families and carers, for schools, and for those who will want to use the estimated grades in decisions about further and higher education and employment. For the RSS, with our Data Manifesto[7], this is partly a matter of principle. However, we think greater transparency earlier might also have improved the adjustment procedure and helped ensure it was more likely to be seen as fair, before individual estimated grades were issued.

As we suggested in our evidence to the House of Commons Education Select Committee, we think it is urgent to have some review of the statistical adjustment procedures used, led perhaps by the UK Statistics Authority's Office for Statistics Regulation. This should consider both the substantive issues of the data used and the adjustment algorithms of the various nations, but also whether greater transparency would have been possible and beneficial.

In the meantime, Universities UK and UCAS may wish to consider how they will treat students whose predicted and estimated grades are far apart, and whether they will consider prior student achievements in addition to the estimated grades in making their decisions.

That is, of course, to look back. There is also however – as we have said in our response to Ofqual's consultation about exams in 2021 – a forward-looking reason for reviewing matters now.

None of us can be certain that Covid-19 will not again disrupt next year's exams, so that a similar situation could arise next year, though we appreciate that it may be better to take steps to reduce teaching disruptions, and particularly differential disruptions in different schools, so that exams can go ahead. There should be a full and open appraisal of what choices have been made this year about the statistical adjustments, whether different data might have allowed more individual-based adjustment, and whether it might have been possible to take uncertainty more formally into account. These are not simply technical statistical choices, but choices about which different people will have different views. Accountability requires that a range of choices are considered and public justification given for the choices underpinning any statistical adjustment procedure used.

Otherwise, we will all be left to grade the exam qualifications authorities' performance as 'could do better'.

Explanatory note

[i] The Royal Statistical Society (RSS) is a learned society, a professional body for statisticians and data analysts, and a charity, which promotes statistics for the public good. We have around 10,000 members in the UK and across the world. Since our foundation in 1834, we have engaged continuously with government, organisations and professionals. We advocate for best practice in the use of statistics and data to enable evidence-based decision-making in the public interest.

[7] Royal Statistical Society Data Manifesto, accessed 5 August 2020, https://rss.org.uk/policy-campaigns/policy/our-data-manifesto/

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

[ii] The Education Policy Advisory Group (EPAG) seeks to influence education curricula and policy to improve the quality of the teaching, learning and assessment of statistics and data. The EPAG provides advice and helps to formulate the Society's consultation responses, research and policy work in statistical education. The EPAG also liaises on behalf of the Society with other key influencing organisations on education research and policy.