

Written evidence submitted by the Royal Statistical Society

Royal Statistical Society response to the House of Commons Education Select Committee call for evidence: The impact of COVID-19 on education and children's services inquiry

8 June 2020

Summary - The effect of cancelling formal exams

1. The [Royal Statistical Society](https://rss.org.uk/)¹ (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.
2. The RSS welcomes the opportunity to respond to the call for evidence of the Committee's [inquiry](https://committees.parliament.uk/work/202/the-impact-of-covid19-on-education-and-childrens-services/)² on the impact of COVID-19 on education and children's services. Our response focusses, in particular, on the statistical issues relating to the cancellation of formal exams. Our response to this call for evidence has been informed by discussions in the RSS Education Policy Advisory Group, together with the expertise of some of our distinguished Fellows in academia.
3. We acknowledge at the outset how difficult the decision must have been to cancel formal exams. As statisticians, we are aware of the social changes brought about by COVID-19; one that is particularly important for younger people and their life chances is about how to estimate grade outcomes for students whose exams cannot be taken within the normal timetable. We are aware too that some (including some RSS fellows) may feel that all students should just have had exams deferred. Such a decision would have brought some advantages in the form of being a known process, yet also some disadvantages, not least in delaying or derailing opportunities at a time when getting back to some sense of normality is an important aspiration for many, and when there has been significant disruption to study regimes.
4. Ofqual intends that exam boards will collect student grade assessments, and student rankings, from teachers within exam centres, and then use statistical processes to obtain estimated grades for students, which will be their GCSE, AS and A Level grades for 2020. To carry out this task, fairly, at this scale, is extremely difficult and unprecedented. Little detail is known about Ofqual's plans. We would welcome more transparency. We are especially concerned about the collection and use of student rankings, as we believe that there is a high degree of uncertainty about these, particularly for students who are not the very highest and lowest ranked.
5. We have some concerns about the lack of clarity and transparency so far provided by Ofqual about their plans for estimating grades.

Engagement with Ofqual

¹ Royal Statistical Society website, accessed 13 May 2020: <https://rss.org.uk/>

² UK Parliament, Education Select Committee, accessed 13 May 2020: <https://committees.parliament.uk/work/202/the-impact-of-covid19-on-education-and-childrens-services/>

6. We wrote to Ofqual on 9th April, after the decision to cancel exams was announced. This was before any Ofqual External Advisory Committee was announced, but we had not been approached at the time it launched its consultation with the details of their advisory group.
7. In that letter, the RSS stressed its keenness to be constructively helpful in extraordinarily difficult times. We appreciated the gravity and urgency of the situation, and the importance of this issue for students, schools, colleges, parents, carers and teachers. We understood that this requires a solution at pace, and what might be ideal if time were less pressed may well not be possible under these circumstances.
8. In both that letter, and our response to the Ofqual consultation, the RSS highlighted a number of statistical issues with the proposed procedures that might be appropriate to consider in 'standardising' centre assessed grades.
9. In the [RSS Data Manifesto](#)³, one of the three core pillars we set out is the importance of using data to strengthen democracy and trustworthiness. We emphasise that democracy relies on the quality and reliability of data in the public domain to ensure that data and statistics are used in a way that people can trust. We are also keen that the statistical processes involved in Ofqual's procedures should adhere to the [UK Statistics Authority's Code of Practice for Statistics, Edition 2.0](#)⁴.
10. We were glad to see from its consultation that Ofqual has convened an external advisory group; those of its members who are known to us are serious and knowledgeable about the issues. We believe that having an engaged external expert advisory group with statistically-informed members is vital as the detailed methodologies both for collecting and evaluating data are put in place. It could reduce the possibility of significant outside challenge over the estimation methods used.
11. The RSS has offered to help by providing other experts who can comment on the grade estimation methodology, and stands by ready to help, either in reviewing plans or more developmentally. We hope the Committee will discuss the statistical issues, and the uncertainty around them, in their engagement with Ofqual.
12. We should stress that we understand that members of such a group should not be giving a running public commentary on the statistical choices (or even the quality of the options), but given the principle of transparency, we believe it might be possible for Ofqual to provide more information about its grade estimation methodology, and the uncertainty that might attach to parts of it. At the early stages, this might have included a general description of the types of analyses it was planning, and the issues it was taking into account. Of course later, we assume they will be even more transparent about statistical adjustments they put in place (without divulging confidential information). This would help inform later, more considered, examination of data options and statistical issues of adjustment after this year's results are published, if only to inform choices that might arise in future.

³ Royal Statistical Society website, RSS Data Education Manifesto, revised edition, 2019, accessed 13 May 2020: <https://rss.org.uk/RSS/media/File-library/Policy/2019/9522-RSS-Data-Manifesto.pdf>

⁴ UK Statistics Authority's Code of Practice for Statistics, Edition 2.0, accessed 8 June 2020: <https://www.statisticsauthority.gov.uk/wp-content/uploads/2018/02/Code-of-Practice-for-Statistics.pdf>

Issues with data collection/ quality

13. In our initial letter to Ofqual, we raised a number of illustrative statistical issues that might stem from the data available to estimate individual student grades. The process of collecting individual student data is now largely complete. One of our main concerns was about how the proposed data collection process would tackle the issue of uncertainty in assigning not only predicted grades to students, but in assessing the degree of 'certainty' in those predictions. (We note of course that even with exams, results are affected by marker variation and the fact that student performance will vary from day to day.) One particular concern was with predictions for middle-ranked students, where we believe uncertainty will be higher than for higher- or lower-ranked students.
14. It is not clear what steps have been taken in data collection to consider the issue of uncertainty, and if they are planning to address the issue of whether certain types of students would be more likely than others to under- or over-perform in an actual exam compared to grades produced by an estimation procedure based on teacher assessment and historical information. We would assume Ofqual is considering this, so in the absence of published details we cannot comment specifically, but there are ways of collecting predictions with uncertainty estimates and/or to consider models that take account of degrees of uncertainty. Alternatively, it would still be possible to use (the individual distribution of) GCSE attainment to illuminate uncertainty, or to use some sort of sampling estimate (based on Centres) of possible uncertainty. Again, Ofqual may well be considering these issues, but the Committee may wish know more about this.

Issues with subsequent statistical modelling of anticipated grades

15. In addition to the statistical issues related to the collection of individual student grade and rank estimates, other statistical issues will arise in the statistical adjustment to produce final estimates. Again, the RSS should emphasise that we do not claim the issues below are exhaustive, and there is insufficient information available for us to take a view on Ofqual's plans or the options available.
16. There are a set of issues linked to the use of evidence about the previous attainment of schools/Centres. There will be uncertainty about this too, especially if schools or colleges have taken steps to improve previous attainment, but also due to natural volatility, and these may be particularly important in the case of exams with a small number of students per school or subject. The Committee may wish to confirm that Ofqual is considering these issues.
17. There are broader questions (some of which may have some prior empirical evidence) about whether there is evidence (either generally, or for specific subjects) about whether particular groups of students tend systematically to exceed or fall beneath their predicted grades/coursework attainment. For instance, we are aware of some evidence that boys tend to achieve higher grades in exams than in coursework and this may affect predicted grades. Some RSS fellows have asked whether there is evidence about bias in predicting results for students from a BAME background. Some of these data may not be available to Ofqual, nor would there be any obvious statistical means of addressing them, since non-statistical judgements are involved. But in addition to these social policy issues, there remain matters to do with considering how to address uncertainty with middle-ranked students, and whether there is evidence from previous years that some teachers, as well as Centres, are more reliable in their predictions of exam results. These too are issues the Committee may wish to explore with Ofqual.

Ofqual proposals for appealing results

18. The RSS in our Ofqual consultation response recognised the challenges of the circumstances posed by COVID-19, and how this could affect the feasibility and desirability of appeals – whether by individuals or centres, on various grounds.
19. It is now some weeks since Ofqual directed exam boards started the process of collecting individual level grades and working on the estimation procedure. It might have been possible, even within the extraordinary time constraints, and the need to protect some types of data confidentiality, for Ofqual to have been more transparent in the methods it planned to use, if only in broad outline. Without that, neither the RSS, nor other external experts, can comment on their robustness.
20. In the absence of greater transparency, we do not believe that the process can be beyond a global appeal as to its fairness or accuracy. One *procedural* way of addressing this would have been to include, as we have suggested, a greater element of external expert statistical engagement with these issues, and so have a robust process of reviewing the statistical methods that makes the denial of appeals more reasonable. Failing this, perhaps a detailed internal review by, for instance, the Office for National Statistics or even the Office for Statistics Regulation (part of the UK Statistics Authority) would be appropriate. We understand that in the time available it will not be possible to consult on the detailed statistical methodology. However, without a stronger procedural basis to ensure statistical rigour, and greater transparency about the issues that Ofqual is examining, it cannot be clear that the statistical methodology will be beyond question.
21. We understand that being transparent about these issues may mean that students whose grades are likely to be more uncertain relative to other students will be more likely to choose to sit deferred exams, now proposed for October and November. Individual students and their families would then have to choose between delaying progress to university and sitting an exam (with all the disruptions to teaching and study schedules since March). However, we believe that transparency means that not only would individuals have more informed options, it would help all of society consider how to fairly acknowledge uncertainty in these unique circumstances.

Additional remarks

22. The RSS still stands ready to help in any way that would be helpful.
23. In any case, we believe that the Committee may wish to consider the issue of how to ensure sufficient external constructive challenge for the exam estimation procedure, while accepting that Ofqual must make final decisions.
24. Finally, we suspect that there need to be further longer-term discussion of the issue of teacher-assigned grades, and how they might be approached in future. It is by no means clear that this year's situation will not arise in future, either because of further waves of COVID-19, or because there are broader questions about the various national examination systems used in the UK (between exam boards, and given different national systems of qualifications), and even whether teacher assessment rather than exams might actually have benefits. The RSS is always happy to engage in constructive discussions, either orally or in writing, of the statistical issues involved.

June 2020