

Education Committee

Oral evidence: [The Impact of Covid-19 on education and children's services, HC 254](#)

Wednesday 2 September 2020

Ordered by the House of Commons to be published on 2 September 2020.

[Watch the meeting](#)

Members present: Robert Halfon (Chair); Apsana Begum; Jonathan Gullis; Tom Hunt; Dr Caroline Johnson; Kim Johnson; David Johnston; Ian Mearns; David Simmonds; Christian Wakeford.

Questions 943-1059

Witnesses

[I](#): Roger Taylor, Chair, Ofqual, Dame Glenys Stacey, Acting Chief Regulator, Ofqual, Dr Michelle Meadows, Executive Director for Strategy and Research, Ofqual, and Julie Swan, Executive Director, General Qualifications, Ofqual.

Written evidence from witnesses:

Examination of witnesses

Witnesses: Roger Taylor, Dame Glenys Stacey, Dr Michelle Meadows and Julie Swan.

Q943 **Chair:** Good morning. Welcome everyone, to the first sitting of the Committee since the parliamentary recess, and welcome to our witnesses from Ofqual. For the benefit of the tape and those listening outside, may I ask the witnesses to introduce themselves very briefly and to give their title? Also, may I just check that you are happy with our using your first names during this sitting? Just nod if you are. Thank you. May we please start with you, Roger?

Roger Taylor: I am Roger Taylor. I am the chair of Ofqual. Mr Chairman, I will just say that if you direct questions to me, I will then, if necessary, ask my colleagues, Julie and Michelle, to address aspects that they would be better able to answer.

Julie Swan: Good morning. I am Julie Swan, executive director of general qualifications at Ofqual.

Dr Meadows: Good morning. I am Michelle Meadows, executive director for strategy, risk and research.

Chair: Dame Glenys, do you want to introduce yourself, although you are not answering questions today?

Dame Glenys Stacey: Good morning. Thank you for that opportunity. My name is Dame Glenys Stacey. I was very recently appointed as acting chief regulator at Ofqual.

Chair: And would you like to be addressed as Dame Glenys or Glenys?

Dame Glenys Stacey: I am happy to be called Glenys or Dame Glenys. I will answer to either.

Chair: Thank you for coming. I wish you well in your new position.

Dame Glenys Stacey: I appreciate that. Thank you.

Q944 **Chair:** May I urge you all to give concise answers? We will mostly direct questions at individuals, or we might say, "Roger, can you answer this?", or, "Can Julie or Michelle answer these questions?" Please be as concise as possible because we have quite a bit to get through today.

We received your statement last night; the Committee will publish it in the next 15 minutes or so. Thank you for sending it. I have to look at the statement very carefully, but my first reading of it suggests that Ofqual is saying that you did everything Ministers asked you to do; you think the algorithm was fair, despite the anomalies; and what went wrong was that pupils, families and teachers would not accept the grades because, in essence, they had not done an exam. So what you are really saying, in



HOUSE OF COMMONS

three words, if I had to describe it in a nutshell, is, “Not me, guv”. Is that a fair reflection of the statement, chair?

Roger Taylor: As we made clear in the statement, we fully accept our share of responsibility for what has gone wrong this year. I have personally apologised to students and parents for what we recognise has been an extremely anxiety-making incident. It has been disruptive to this year’s candidates, disruptive to higher education, and disruptive to next year’s candidates.

Q945 **Chair:** Thank you. You say that in your statement, but, specifically, is my summary of the statement accurate?

Roger Taylor: What we would point towards is that the fundamental mistake was to believe that this would ever be acceptable to the public. Perhaps I might illustrate that point to give a sense of what I mean. We were looking at this in terms of particular ideas of fairness across a whole population and the probability that somebody would get a grade, but acknowledging from the outset that it would not be anything like as accurate as exams. What we now realise is that if you have 1,000 students that have, for example, an 80% chance of getting an A grade, they would regard themselves quite reasonably as A-grade students. What we were doing in effect was recognising that, in a normal year, 200 of those students would fail to get their A grade. They would go into the exam and it just would not happen on the day, and they would get a B grade. We were using statistics and teacher ranking in order to—

Q946 **Chair:** Okay. In the statement, you said there was no evidence of bias, but you knew that larger institutions such as sixth form and FE colleges were disadvantaged by the algorithm, and those institutions are more likely to educate those from lower socioeconomic backgrounds. Is it not extraordinary for your statement to say, “We knew about this, but were unable to find a solution to this problem”? If you knew about it, why didn’t you ask for help?

Roger Taylor: We did know about it. It is important to say that given the choice between standardising and not standardising, the impact is better for students from lower socioeconomic statuses with the standardised results. Despite the known issues with small centres—again, the same is true of private schools, as the inability to standardise small cohorts gave an advantage to private schools—it is also true that, overall, the process of standardisation reduced the advantage enjoyed by private schools. That is why we felt that it was fairer to use the standardisation process as a mechanism to ensure the greatest possible fairness in the circumstances. We acknowledge that the level of fairness achieved was not felt to be acceptable, but it did improve the level of fairness.

Q947 **Chair:** If I could turn to the relationship between the Department for Education and Ofqual, we know that, on 31 March, the Secretary of State issued a direction asking Ofqual to mandate the method of calculating final grades and to develop an appeal process for GCSEs, AS and A-levels, as well as vocational qualifications. On 15 August 2020, Ofqual



HOUSE OF COMMONS

published criteria by which mock exam grades could be used in the appeal process and then subsequently retracted. The DFE had also said that the distribution of grades should follow a similar profile to that in previous years. On 19 August, a DFE statement implied that Ofqual took the decision to allow pupils to use centre assessed grades, and the decision was “one that the Department agreed with”. What these events show is that perhaps there is not just one clear authority making the final decision. Could I ask you, Roger, to clarify in simple terms the official legal relationship between the DFE and Ofqual, and what authority each institution had for which of the major decisions that were taken?

Roger Taylor: Certainly. The relationship is one in which the Secretary of State, as the democratically accountable politician, decides policy. Ofqual’s role is to have regard to policy and to implement policy, but within the constraints laid down by the statute that established Ofqual. Those constraints are that the awarding of grades must be valid, it must maintain standards year on year, and it must command public confidence. We can decide not to implement a direction from the Secretary of State if we feel that it would directly contradict those statutory duties, but if the policy does not directly contradict those statutory duties, our obligation is to implement policy as directed by the Secretary of State.

Q948 **Chair:** So who had the authority for which decisions when it came to the development of the standardised model algorithm and which grades should be used? At what point, precisely, was the decision made on 15 August—that Saturday—to retract the statement on mock exam results in appeals, and who took that decision?

Roger Taylor: Perhaps I should walk it through from the beginning, because there is an important relationship in terms of the other duty Ofqual has: to offer advice to Ministers. At the outset, our initial advice to the Secretary of State was that the best way to handle this was to try to hold exams in a socially distanced manner. Our second option was to delay exams, but the third option, if neither of these was acceptable, would be to try to look at some form of calculated grade. We also looked at whether that might be a teacher certificate, rather than attempting to replicate exam grades. That was our advice to Ministers. It was the Secretary of State who then subsequently took the decision and announced, without further consultation with Ofqual, that exams were to be cancelled and a system of calculated grades was to be implemented. We then received a direction from the Secretary of State setting out what he wished Ofqual to implement. At that point it became our responsibility to decide how to implement this in terms of the design of the algorithm that was used to standardise results in line with the direction of the Secretary of State.

Q949 **Chair:** You took the decisions entirely when it came to the development of the algorithm.

Roger Taylor: That is correct.

Q950 **Chair:** In terms of 15 August and the retraction of the statement, you put out a statement in the afternoon on the Saturday, and by late Saturday night it had gone. Who made that decision?



Roger Taylor: To walk you through the decisions there, the Secretary of State informed us that, effectively, they were going to change policy. Until that point, the policy had been calculated grades plus an appeals process. The Secretary of State informed me that they were planning to change this policy in a significant way by allowing an entirely new mechanism by which a grade could be awarded through a mock exams appeal. Our advice to the Secretary of State at this point was that we could not be confident that this could be delivered within the statutory duties of Ofqual, to ensure that valid and trustworthy grades were being issued. The Secretary of State, as he is entitled to do, none the less announced that that was the policy of the Government.

That having been announced as the policy of the Government, the Ofqual board felt—I think correctly—that we should therefore attempt to find a way to implement this in a way that was consistent with our statutory duties. We consulted very rapidly with exam boards and other key stakeholders. We were very concerned that this idea of a valid mock exam had no real credible meaning, but we consulted very rapidly and developed an approach that we felt would be consistent with awarding valid qualifications. We then agreed that with the Department for Education and, to our understanding, with the Secretary of State's office. We then published this on the Saturday. We were subsequently contacted by the Secretary of State later that evening and were informed that this was in fact not, to his mind, in line with Government policy.

Q951 **Chair:** At what time were you contacted?

Roger Taylor: It was published about 3 o'clock on the Saturday. I think the call from the Secretary of State was probably at around 7 o'clock, 8 o'clock that evening. The Secretary of State first phoned the chief regulator. The chief regulator explained that this decision had been made by the board and had been made on what I think were extremely sound principles, and that therefore she could not, as it were, take down the statement or reverse the policy.

The Secretary of State telephoned me and said that he would like the board to reconsider. I again reiterated what struck me as the very clear arguments as to why this was the right way to implement this policy and that to do otherwise would have been inconsistent with awarding of valid and reliable grades. None the less, given the Secretary of State's views, it felt appropriate to call the board together very late that evening. The board convened at, I think, around 10 o'clock that evening.

I think at this stage we realised that we were in a situation which was rapidly getting out of control—that there were policies being recommended and strongly advocated by the Secretary of State that we felt would not be consistent with our legal duties, and that there was, additionally, a growing risk around delivering any form of mock appeals results in a way that would be acceptable as a reasonable way to award grades, the issue here being simply that this was looking increasingly like it was going to result in 85%, 90%, of candidates receiving a centre assessed grade in effect, but potentially a small number not receiving them, because they



HOUSE OF COMMONS

did not have access to this route. Furthermore, it would take time to make this happen and therefore there was a risk to candidates' university places; and there was, furthermore, an increasing risk about the deliverability of it. So for these reasons we felt that we were now in a situation that was moving rapidly out of control and that it was likely that the only way out of it was to move towards centre assessed grades.

Q952 **Chair:** But when you put it up on your website, did you not agree it with the Department first, before you put that information up?

Roger Taylor: Yes—and with the Secretary of State's office.

Q953 **Chair:** So before the information went up on the website on Saturday 15 August in the afternoon, you agreed that. You put it up on your website, and what you are saying is that the Secretary of State called you and asked you to change that information.

Roger Taylor: Yes.

Q954 **Chair:** Why did you then put on the website, "we have taken the information down, and the board will be considering options"? Why didn't you give a fuller explanation of what was going on?

Roger Taylor: Because we felt we needed to reconvene and discuss further with the Secretary of State the options that were available. The three options that were then available were to implement the mock appeals route as the Secretary of State envisaged it; to go back to the situation that we had originally proposed, which was a mock appeals route that was consistent with centre assessed grades; or the third option was simply to move to centre assessed grades. At this point it was clear that to have given any clearer guidance without having had an opportunity to discuss these issues further with the Secretary of State, and tried to move towards what we felt was a policy that would be fairest to students and most likely to deliver results in good time and order—

Q955 **Chair:** Right, so by a couple of days later the decision was made for the U-turn, to accept the centre assessed grades.

Roger Taylor: There were discussions throughout the Sunday, and on the Sunday evening Ofqual took the decision that there was no way to pursue this in any orderly, reasonable manner that would deliver fair grades for students except through—

Q956 **Chair:** So on Saturday the 15th, the appeals system was going to be widened to include mocks, which we know because it was put on your website. The Secretary of State then calls and says that you cannot do that, so by Sunday, 24 hours later, the move was made to centre assessed grades. Is that what you are saying? Is that correct?

Roger Taylor: That is right.

Q957 **Chair:** I will bring in my colleagues soon, but can I just ask whether Ofqual challenged the Department for Education on any decisions made by the Department during the development of the algorithm or its request to avoid grade inflation and have it like previous years and so



HOUSE OF COMMONS

on?

Roger Taylor: We initially advised against cancelling exams. With regard to the maintenance of standards and the design of the algorithm, that was a process of continuous collaboration. We felt that this was not like other years. It was not that we were regulating the independent awarding of grades by exam boards; we were effectively constructing, together with the Department and the exam boards, a national system to determine how grades were to be awarded. It was therefore not appropriate to do anything except involve them very closely in all decision making. There was a programme board that oversaw this, which involved the exam boards, Ofqual and the Department for Education.

Chair: Before I carry on, I will bring in Tom Hunt and Jonathan Gullis very quickly, and then I will carry on and then bring in my other colleagues, if they have some quick, relevant questions.

Q958 **Tom Hunt:** Just a point about how, when Ofqual takes positions, it takes into account whether it can secure broad public support. Can you understand why some of us find it odd, strange and hard to believe how anyone could think that something that would so penalise state schools compared with private schools, with a rule that class sizes below 15 had greater weight placed on the CAGs than larger classes. That fed through into a significant increase in the number of A*s and As in private schools but not in comprehensive schools. Do you agree with us that it is hard to really understand how anyone could ever think that that would secure broad public support?

Chair: Just a brief answer, please.

Roger Taylor: I would like to reiterate that the standardisation process reduced the advantage enjoyed by private schools. It is important to say that. On your broader point, from the very outset we said that this is an enormously difficult thing to do in a way that will command public support. However, we have to have due regard for Government policy. If a Minister decides that this is the policy that the Government wish to pursue, there are very good reasons for this policy. It ensures greater fairness in the sense that you have less variation in grades between centres. It ensures greater fairness in that next year's candidates will not be affected by these results. It ensures greater fairness in the sense that it results in less disadvantage for disadvantaged communities—

Chair: Thank you. Jonathan Gullis, briefly, please.

Q959 **Jonathan Gullis:** I would like to understand what Ofqual did once it had seen what happened in Scotland regarding standardisation and the algorithm. Did you not take into account what happened in Scotland and therefore make adjustments?

Roger Taylor: Perhaps two comments about this. First, the initial response in Scotland was driven by a view that it had unfairly penalised disadvantaged communities. As I have said, we were confident that our approach was not unfairly penalising disadvantaged communities.



HOUSE OF COMMONS

Secondly, the decision to change policy, on the basis that there was no public confidence, was obviously something that Ministers would need to agree with. It was on that day that the Secretary of State decided that the Government's response, in terms of the policy adjustment in relation to Scotland, would be to introduce this new form of appeal based on mock grades. Our advice was that that was extremely risky, because we were unclear as to how it could be delivered in a way that was consistent with Ofqual's legal duties.

Q960 David Johnston: There are various aspects of this that have the feel of a slow-moving car crash. One is about high-attaining children at low-performing schools. You identified that that would be an issue, and you thought it affected about 0.2% of young people, but you said it was too difficult to identify them. Did you consider asking schools to identify them, because they are very used to doing so for university references and so on?

Even if you could have resolved all these issues in the appeals process—even if you could resolve every single one of those cases, which itself is questionable—you must have realised the distress it would cause on results day.

Roger Taylor: Thank you for the question. Yes, as you can imagine, it was a question that absorbed a lot of our time and thinking. We had run a number of mechanisms for identifying candidates, and one of the mechanisms was indeed to use the information from the centre assessed grades in the sense that a school that, on the whole, did not present very inflated grades but had one student performing very differently from the rest of the cohort was a signal that this was indeed occurring.

It was clear that to make a valid judgment would require a degree of human judgment and therefore a form of appeal would be necessary to make this work, but we were also exploring with the exam boards how we could implement a system of outreach to those students through the exam boards to let them know on the day, "Look, we think you've probably got a very good case for appeal." That was the direction we were moving in. When the mock appeals route came in, that question became less relevant.

Q961 Chair: Could I challenge you on the situation in terms of disadvantaged students? You say in your letter that your model did not disadvantage disadvantaged students. The Department for Education confirmed on 14 August that pupils from lower socioeconomic groups were more likely than their peers to have their centre assessed grades downgraded by Ofqual's algorithm at grades C and above. The difference between Ofqual's moderated grades and teacher centre assessed grades for lower socioeconomic groups was 10.42%. In contrast, the difference between Ofqual's moderated grades and teacher centre assessed grades for higher socioeconomic groups was 8.34%. Did Ofqual identify these as potential issues at any point? Did you raise these issues with Ministers? What was their response? I am happy if one of the other witnesses answers, if you feel that is necessary.



HOUSE OF COMMONS

Roger Taylor: I might hand over to Michelle here, but I will just say that a correct interpretation of the figures is that the moderation process gave strong candidates from disadvantaged communities a better relative outcome than the unmoderated results. I will hand over to Michelle now.

Q962 **Chair:** Okay. Just before you answer, Michelle, we know that private schools benefited. Your report on the algorithm published after the results were issued mentions the exclusion of “data with respect to Learners registered with independent or selective Centres.” What was sought to be achieved by this exclusion?

Dr Meadows: I will take that question first and then come back to the question about the changes to centre assessment grades. We used the national pupil database as our source of data. Often for certain centre types—independent schools—the data is patchier. We can come back with a full detailed response to your question, but it will be that that data was not available for that particular centre type. We will come back with the full technical detail for you.

On the question of the changes to the centre assessment grades by socioeconomic status, of course what really counts is the final grade that a student walks away with—that is what actually impacts on their life chances. We had done a full equalities analysis, looking at the grades not just by socioeconomic status but by other protected characteristics such as ethnicity, gender and so on, and what we were able to see and we were very confident about was that any fluctuation in outcomes seen for these various groups this year was extremely similar to the small changes in outcomes we had seen in previous years. In other words, there was nothing about the process that was biased.

Q963 **Chair:** As I understand it, you developed 11 algorithms and tested them using 2019 data. Is that right?

Dr Meadows: There are 11 or 12 algorithms in total that were tested.

Q964 **Chair:** Schools and colleges were asked to submit their centre assessed grades by 12 June 2020. Is that correct?

Dr Meadows: I believe so. I would have to check the exact date.

Q965 **Chair:** Why then did Ofqual not use the time between 12 June 2020 and the 13 August A-level results to test the algorithms using a sample of data from 2020 and compare it with the centre assessed grades, rather than wait until the eleventh hour to realise the algorithm produced discrepancies? Some of the wild anomalies that sixth form colleges saw would surely have become apparent much sooner. In essence, I am asking whether you should have done your own mock exam in terms of the algorithm.

Dr Meadows: We tested the model thoroughly. We tested 11 or 12 different approaches in total. We tested the impact on different centre types and centres with different proportions of candidates from different ethnicities, socioeconomic statuses and so on. We were confident that the



HOUSE OF COMMONS

model we chose was the most accurate overall and the most accurate for those different groups of students. We did not do this in isolation.

Q966 **Chair:** No, but did you do a dummy run, in essence, to check out what the anomalies were for all students before the final version of it? If you had done that, you might have been able to pick out and sort out the problems that then came about.

Roger Taylor: Perhaps I could jump in here. We were obviously aware of this and throughout June were looking at those issues. Indeed, there were ongoing discussions about different ways of defining what an anomaly was. There were two issues that we were considering during that period, one of which was whether there was a mechanism by which we could apply or overlay additional rules to address these in advance of a warning. It was very clear, and the legal advice on this was extremely clear, that that would involve us making arbitrary judgments that would not be justifiable.

Q967 **Chair:** Did you raise the predictive accuracy of the model with Ministers? Were the risks of the model raised with the Department for Education?

Roger Taylor: Throughout the process, yes. I might hand over to Julie to say a little bit more about that. I just want to make one further point: the mechanism that we realised was the best way to deal with these issues that you are talking about was the appeals process, and that led to the guidance on appeals that was issued, which made clear that even an individual outlying candidate would be able to make an appeal on the grounds that standardisation had adversely affected them.

Q968 **Chair:** Who was on your external advisory group, and how was it appointed? Did you ask experts and external representatives to ensure the model was fair? Could one of you answer that as well briefly please?

Roger Taylor: If Michelle answers that one, perhaps Julie could come back to your earlier question about informing the Department of the risks. Michelle, do you want to address the question about the external advice?

Dr Meadows: We had 10 members of the external advisory group. Only one of those members worked within an organisation associated with the exam boards—somebody from Cambridge Assessment—and the rest were from external organisations including Ofsted, UCAS, the University of Oxford, independent consultants and so on. They were chosen to have a blend of statistical and assessment expertise.

Q969 **Chair:** But not the Royal Statistical Society?

Dr Meadows: We did have a member who was associated with the Royal Statistical Society—indeed, he is a member of their education policy group.

Q970 **Chair:** As you know, they are not very happy, because they feel that they offered a lot of help and you imposed such stringent conditions that it was impossible for them to do so. To promote transparency, will you publish all the communications and minutes you have had with the



HOUSE OF COMMONS

Department?

Roger Taylor: Yes, I think we can do that. We obviously need to discuss with the Department whether there is any form of deliberative privilege that they would wish to execute. We certainly do not have anything to hide, but I don't think it is appropriate for us to simply say we will publish correspondence with them without that discussion.

Q971 **Chair:** Finally, before I bring in my colleagues, apart from the issue you have talked about on the Saturday night, what discussions and meetings did you have with the Department between A-level results day on 13 August and the Government's announcement that centre assessed grades would stand on 17 August?

Roger Taylor: Julie, I might hand over to you to refer to the full record, but we had conversations on the day before—both Sally and I, I think, had conversations; I certainly had a conversation with the Secretary of State and a number of conversations with policy advisers to the Secretary of State. We fed back our concerns about the proposal and issues that we felt would be necessary to make it implementable—

Q972 **Chair:** And what did the Secretary of State say on the Sunday to you, having defended the system—the Government defended the system for the previous days before, on 16 August—that he was accepting that there would be a U-turn? Is that right?

Roger Taylor: We had a number of conversations on that day. I think by the end of the day he had come to accept and recognise that Ofqual—that the situation was now one where any other policy would simply be indefensible, and that we had to make that decision. He did say, if we made that decision, we would work collaboratively to try to implement it.

Q973 **Dr Johnson:** We know that 75% of young people taking the exams would not have, under ordinary circumstances, achieved the grade that they were predicted in some subjects. Do you think, on that basis, therefore, that the mistake was in cancelling exams? Once the exams had been cancelled, realistically, any statistical process that you made was always going to leave people disappointed and not meet your requirement of broad public support.

Roger Taylor: I think that is exactly right. We did say at the outset that the risk is that statistical prediction is simply not accurate enough, and that there would definitely be a candidate who would know that they would have done better in the exam, but who would be given a result lower than what they reasonably and competently felt that they could do in an exam.

When the decision was originally made, there was a strong belief that the autumn series would be the compensation for that—that people would be given a chance and that university places could be held open for them that they could take in January, and that that would limit that damage. At the time, it was felt that it was a fair offer, but of course, over time, schools did not reopen; there were no arrangements for late entry to university; and by July, it was clear that the autumn series did not represent any sort



HOUSE OF COMMONS

of reasonable alternative that candidates felt would make up for being given an inaccurate calculated grade. At that point, we were in a situation where it was difficult to see how people would accept it as a fair way to have their grades awarded.

Q974 Ian Mearns: Roger, you started off this morning by saying that you have already apologised and that you apologise again but, I am afraid to say that, given the huge amount of information and really dreadful stories that we have heard directly from students, their parents and their centres, in this context, an apology really doesn't cut the mustard.

Michelle, in answering an earlier question about how we got here, you said that you had tested the model thoroughly, but the trouble is that you tested the moral thoroughly and ended up here, and you couldn't see that you were going to end up here, and what the ramifications of ending up here were going to be. The fact that you could not see that, from my perspective, given the fact that you are a public body overseeing this sort of process, is a huge problem—that you couldn't see that there would be a huge public backlash to this whole scenario.

Therefore, from that perspective, I have to ask whether the prime directive in starting the whole process was to avoid what these days is called grade inflation, so that the process would be seen to be consistent with previous years' results, because I have a problem with that premise. This is the problem: Ministers are regularly telling us that we have more good and outstanding schools, with the most highly professional teaching profession that we have ever had. Given that process, that improvement and that continuing improvement, should there not be some increase in the levels of achievement by youngsters year on year that cannot be put down as grade inflation?

Roger Taylor: That is a very good question. There are two issues there. In terms of the awareness of the difficulty of landing this mechanism in a way that the public would have confidence in, it is important to stress that Ofqual's duty is to attempt to implement Government policy. There were very good reasons why standardised grades were preferable to unstandardised grades. I do not think that it would have been appropriate for Ofqual to have ignored the Secretary of State's guidance on the basis that we did not believe that it would ever be deliverable in a way that the public would trust. That would have been going beyond what the reasonable behaviour of an independent regulator would be.

On your point about grade inflation, we were very aware that being very strict about grade inflation would only make this situation worse. That is why, in the design of the model, at every point where we could reasonably do this, we erred in the direction of making decisions that allowed grades to rise. Consequently, the final result of the moderated grades did allow for between 2% and 3% inflation in grades which, in assessment terms, is very significant and larger than would represent the sorts of effects that you talked about resulting from improvements in teaching, but we felt that that was appropriate in these extremely unusual circumstances, given the disruption happening in people's lives as a result of the pandemic.



Q975 **Ian Mearns:** But Roger, I have also seen, importantly, some suggestion that if, on an annual basis, all the scripts that were marked—6 million or so at GCSE and A-level—were marked again by a senior examiner, up to 25% of them might have their grades improved. One in four might have their grades improved, and yet the number of appeals is relatively low, compared with that. What are your comments on that? I believe that to be the case.

Roger Taylor: This is a comment on research that Ofqual has published. It is important to understand that no form of assessment is so accurate as to genuinely capture perfectly the abilities of each individual. What you are referring to is the research about accuracy that Ofqual has been leading on. In this context, it is also important to say that the inaccuracy works both ways. But, Michelle, perhaps I can hand over to you, because this is work that you have been doing.

Chair: Briefly, please.

Dr Meadows: Every year, we publish marking consistency metrics that report the extent to which grades would change if a different senior examiner had looked at the work. In fact, we looked at that work this year and took some comfort from it, in the sense that the levels of accuracy that we were seeing from the standardisation model were very similar to those levels of accuracy that we see each year through the marking process.

Q976 **Ian Mearns:** One last comment before we move on is that Ofqual is independent and should maintain its independence, but it seems as though the hand of the DFE and Ministers has been too closely and intricately involved in this whole process. They should have been kept out of it, if Ofqual were to sustain and maintain its independence in the process.

Roger Taylor: I would say that we were in completely uncharted territory. Nobody has ever had to deal with a pandemic situation. As some people have said, this has been the biggest challenge for the country since the second world war. In determining what is the acceptable way forward, I think it is appropriate that that decision falls to democratically elected politicians; I do not think that it would have been appropriate for Ofqual to have determined the fundamental policy of how we were going to proceed.

Q977 **Chair:** But basically what you told us earlier was that, on Saturday 15 August, you were worried about widening the appeals system in the way that the Government suggested, but you still put on your website that you agreed with them. Then the Government ring you up late at night to tell you, "No, this isn't acceptable," and you do what they say and take it off.

Roger Taylor: No. I perhaps have not explained that as clearly as I should have. We implemented the policy in a way that we felt was acceptable. The crucial thing that we had realised was that there was the idea that there was a valid mock result—yes—but what became very clear was that there really were not large numbers of students who had taken,



say, every paper of an examination under examination conditions, and had them marked according to the mark scheme and graded appropriately, which would have been, as it were, pretty strong evidence of the grade that they would have got. What people tended to have was partial evidence from various tests taken at various times.

We had asked the head of centre to sign off—taking into account the existing evidence from mock results—the grade that they thought the candidate was most likely to achieve. If we had then allowed the same centres to submit a partial piece of evidence and say that they believed, on the basis of that evidence, that the candidate would in fact have achieved a higher grade, we would be inviting them to make two directly contradictory statements. Those statements were signed off by the head of centre, and they are extremely serious statements about the reliability of the information that the centres were providing. It would have been completely irresponsible to have allowed that to occur. Therefore, the only way to deliver this policy sensibly was to allow the centres to present—

Q978 **Chair:** But what I am saying is that you still took it off your website when the Secretary of State asked you to.

Roger Taylor: Yes, because the Secretary of State was—

Q979 **Chair:** The independence of Ofqual therefore comes into question, because you could have said, “No, I’m not taking it off the website.”

Roger Taylor: It is important, in trying to manage public confidence, that we do not have a Secretary of State stating one policy and Ofqual stating a different policy. It also struck us that the way to resolve this was to move at pace and it needed to be negotiated and managed in an orderly fashion. But we were acting with full independence.

Chair: I am going to bring in David Simmonds, and then Tom Hunt wants to ask a brief question.

Q980 **David Simmonds:** Thank you, Chair. I am one of those Members whose inbox now has a lot more complaints about the consequences of the U-turn than about the original results that were published in respect of A-levels. The Committee spent a lot of time looking at the fairness of exam results and the process that would be followed. There are a couple of linked questions that I would like to put to our witnesses.

The first builds on the question that Ian asked about the role of Ministers. I would like to hear a bit a more about the degree to which assurances were sought about the risks of the alternative options, which had been highlighted by a number of people, and how effectively those were considered in arriving at a decision about the best way to proceed. I do not think it has been helpful to have a blame game process at the moment, but it is helpful to understand how effective that process of assurance was to learn lessons about avoiding this kind of thing in future.

Secondly, to touch on something that a number of people have mentioned around the appeals process, which I know we will come to in more detail later, does the panel feel that an appeals process in which individual



HOUSE OF COMMONS

students do not have a right of appeal can ever be fit for purpose? It seems to me that the risk is that it places institutional interest ahead of the student's interest, which is certainly reflected in a lot of the complaints that I am getting at the moment.

I have a third and separate question—if it is okay, Chair, I will ask these ones together. Does the panel believe that reliance on final exams is a significant weakness in the education system at the moment? Would it be better to have more alternative methods of assessment? Has this situation demonstrated a lack of resilience in our exams system?

Finally, does the panel have any comment to make about what guidance was provided to schools in putting together centre assessed grades? We have heard a lot of anecdotal comment from teachers, including one who I have spent a lot of time talking to about this, who was told, "Your centre assessed grades will have to be statistically in line with previous years' results, so even if you think you've got three A* students in your given cohort, if you got one A* last year, only one of them is allowed to have an A*." I have not been able to find that reflected in any guidance that was provided to schools, but that seems to be quite a common issue cited by headteachers and teachers as an expectation that they felt they were under, which is clearly an issue that goes to the heart of fairness.

If you could answer those questions on assurance, the point about the appeals process—the fairness and the lack of student right to appeal—the reliance on final exams versus other methods of assessment, and then the point about guidance.

Roger Taylor: I will take the first three. I will hand over to Julie on the last question particularly, but I will give you initial answers.

In terms of the assurance around the original decision, we put our advice into the Department of Health. Our advice was to try and run the exams in a socially distanced way; secondly, delay the exams; or the worst-case scenario is these calculated grades. If we are going down that route, it would be worth considering whether a teacher certificate, rather than attempting to replicate grades through a system that really was not a replication of that system, should also be considered.

Obviously, the decision then is a political decision, and you would need to talk to departmental officials and Ministers to understand exactly the process that was gone through. An announcement was then made two days later, saying that exams were going to be cancelled, and the process as we know it was launched, but I cannot talk in any more detail specifically about how that decision was made.

Q981 **David Simmonds:** May I explore that a little bit more? You said that essentially, your advice was that calculated grades were the worst-case scenario. Did Ministers or Government come back to you and say, "Let's understand the risks with calculated grades, versus the deferral of examinations," versus the other options that might have been out there?



Roger Taylor: I might refer to Michelle and Julie, just to talk about the discussions. The paper that was presented at the time does list in very vivid detail the risks that that approach presented, but of course, it is also important to recognise the risks of trying to implement exams. Quite reasonably, a lot of politicians were very aware of the fact that parents were quite likely—countries that have tried to do this, such as Germany and India, have encountered strong resistance to the idea that it is fair to try and still hold exams. It is important to recognise there were no easy choices; there were no easy options. We were in a pandemic situation, and it was new territory for everybody.

Coming back to the appeals issue, once you have moved to a system where there is effectively no independent, standardised piece of work from the student that is the key point of reference for identifying their skills and knowledge, we are in a situation that changes the relationship with the appeals, because what is it that you are appealing about? I think we were clear that if you are going to implement a situation of this sort, it does not make sense for these students to be able to appeal the teacher's judgment as to the grade they would have got, except in the circumstance where they feel the teacher has basically failed to do the job properly: that the teacher has been biased, or has produced a simply unreasonable and unjustifiable grade. In that circumstance, the student does have a right of appeal, but the notion that you could say, "I think you've got it wrong. I think I would have been a B, rather than a C"—I do not think there is any sensible process by which such an appeal could have been heard. However, that goes right to the heart of why this whole process ultimately feels unfair to the individual, because they have not had the appropriate degree of agency. They have not been able to take some action.

Q982 **Chair:** It could have been signed off by the headteacher. If a student appealed and it was signed off by the headteacher, surely that would have been a very just way of going forward.

Roger Taylor: The headteacher had already signed off the—

Chair: No, no. You were saying, "You can't have a situation where everyone is allowed to appeal because they've got a D instead of a C", or whatever, but if you had a system whereby an individual student wanted to appeal, provided it was signed off by the headteacher, that surely should work.

Roger Taylor: If the student appeals and says, "We think you've got it wrong", the school can simply contact the exam board and say, "I'm sorry, we've made a mistake. We gave you the wrong information." That would not require an appeal; under the mechanism, the school could simply say, "Yeah, okay, we've got that wrong", and inform the exam board.

Q983 **David Simmonds:** If it helps, Chair, one of the things that concerns me about this appeals situation is that one of the reasons why predicted grades were not used previously as part of assessment was the recognition that 75% of the time, they were wrong. They were sometimes better and sometimes worse: disadvantaged students tended



HOUSE OF COMMONS

to be predicted worse than they would actually get, and nice middle-class children tended to get more optimistic grades predicted for them.

I appreciate that centre assessed grades are different from that, but it seems to me that by saying “only the institution has the right of appeal”, you are effectively saying to the institution that it needs to mark its own homework. A lot of the complaints I am getting are from students who are saying that the school is effectively saying to them, “Well, we’ve given you this grade. It’s just tough. We’re not willing to have a further conversation with you about whether your evidence suggests you would have got a different outcome.” The fact that those students in that situation are left, effectively, with nowhere to go, seems to me inherently unjust.

Chair: Can you answer that in a nutshell, chair?

Roger Taylor: It goes to the nature of the problem: there is not an independent piece of information that can be used to determine between these two competing claims. That is why the lack of any form of standardised test or examination makes this a situation that people find very hard to tolerate.

Q984 **Chair:** You wanted to pass to your colleagues the other questions that David asked.

Roger Taylor: I just want to come back to the question of reliance on final exams. It is worth reflecting that the situation is such that other places have also struggled to implement a system of calculated grades. For example, all four territories of the United Kingdom have different arrangements for final exams, but none of them has been able to develop a system of calculated grades that commanded public confidence, so I would resist that conclusion. Julie, you might want to talk about the guidance provided to schools and I think perhaps you wanted to comment on the statement just made.

Julie Swan: I will pick up on a couple of those points. Earlier, I was going to explain some of the challenge we had had from Ministers throughout. We were meeting the Schools Minister on a weekly basis throughout. When we first gave our advice on the options on 16 March, which was written for Ministers, we said it would be challenging, if not impossible, to attempt to moderate estimates in a way that is fair for all this year’s students. Everyone, throughout the process, was aware of the risks. A paper to the general public sector ministerial implementation group on 1 May, highlighted the risk of widespread dissatisfaction with grades awarded, from individual students, schools and colleges, and the risk to public confidence.

We briefed No. 10 on 7 August. The paper written there was very alert to the risks of disadvantage to outlier students, to centres that had expected improved grades this year and to the impact on low-entry cohorts, including independent schools. We had been briefing and we were challenged very clearly throughout on the risks. The lack of alternative has been the problem throughout.



HOUSE OF COMMONS

In terms of the appeals, as Roger said, the real issue is that a judgment had to be made on what grade a student would have been most likely to achieve if they had taken their exams. We set out guidance for teachers to follow. That extensive guidance was updated as questions came to us. Heads of centre signed off the declaration as they submitted their CAGs saying that, in their judgment, this was the grade that the student was most likely to have achieved had the exams taken place.

It was clear to us from the start that in the absence of exams, many students would likely think that they would have got a better grade had exams taken place. Many students might have got a better grade, and many might have got a worse grade than their school or college predicted. The difficulty is that we will never know, because, unfortunately, exams had to be cancelled.

Chair: Julie, if you want to continue, please do. I hope that David can hear what you are saying.

Julie Swan: I just wanted to pick up on the final suggestion—we have heard this too—that some schools interpreted the guidance as suggesting that if they hadn't previously had a student who had received an A* for a subject, they were not allowed to submit a CAG that was an A*. That was certainly not in the guidance at all.

We set out the range of evidence that we asked schools and colleges to take into account, including students' performance in mocks, classwork, homework and any non-exam assessment. We said that to help them check the extent to which their predictions were realistic, they might look back at the prior performance and attainment of their students and adjust accordingly. But there was absolutely no prohibition on a centre putting in a grade for a student simply because a student in a previous year had not achieved that. The declaration that the head of centre had to make was clear that this was about their judgment of the grade each student would have been most likely to achieve, had the exams taken place.

Q985 **David Simmonds:** So we are absolutely clear, I think, Chairman, that any decision that the grades awarded were simply based on reflecting the prior attainment in that institution was taken at the school; it was not a reflection of any guidance that was provided to schools by Ofqual.

Julie Swan: It was included in the guidance to help a centre just to sense-check its centre assessment grades; that was something that they would likely want to take into account. And of course we said that it would be taken into account in standardisation. But what we asked centres to do was to give their judgment, and it's not easy, on what grade a student would most likely have achieved if the exams had taken place.

Q986 **Ian Mearns:** I think there are some concerns over the absolute clarity of the Ofqual guidance to schools on centre assessed grades. It is clear that the guidance, whether from Ofqual's perspective or not, gave the impression to headteachers and their staff that they had to be very, very careful and moderate about the grades that they would award to their



individual students.

Just for absolute clarity, will Ofqual and the board please release forthwith all school submissions for A-level, AS and GCSEs to a trusted third party, such as the Royal Statistical Society, so that a deep forensic analysis can be undertaken to determine how many schools were seeking to game the system, how many were over-optimistic, how many just got it wrong because of Ofqual's guidance, and how many submitted centre assessed grades that were closely in line with the no-grade-inflation guidance?

Unfortunately, all these things seem to have had an impact, not just on centres but on many, many individual students across centres throughout the country. Therefore, from that perspective, those individual students who really are in very difficult situations need some answers on that. So we are looking for reliable, third-party forensic analysis of everything that has been done, because it is important.

Q987 Chair: Okay. Can we have a brief answer—are you going to publish it?

Roger Taylor: We are not going to publish it obviously, because it involves a huge amount of confidential information, but we absolutely and totally agree with what Ian has just suggested, namely that it is absolutely essential that we find—there is an enormous amount that we can learn from what has happened over the summer. And that information is contained in the data submitted by schools and the process of standardisation. It is absolutely essential that independent researchers have access to that in a secure way that will enable those lessons to be learned.

Q988 Chair: Previously, you said you would publish the minutes of correspondence and meetings with Ministers. Will you give us a date when that will happen?

Roger Taylor: I will need to talk to the Department. Obviously, the minutes of ministerial meetings are recorded by the Department. So I think I will need to write back to you on that—

Q989 Chair: Okay. But you will have your own minutes and papers that you can publish. Just to confirm, you will publish those for us to see?

Roger Taylor: Yes, we will. As I say, I will talk to the Department about—

Chair: Okay. That is fine.

Q990 Ian Mearns: Chair, my very last question is about the construction of the algorithms. Who was it who put the guidance in to construct the algorithms in the first place? I understand that a dozen algorithms were actually constructed, and you came up with the one that has ended us up here. So who actually consulted the guidance for the algorithm, who constructed the algorithm and was it outsourced, in terms of constructing the algorithm?

Chair: Who can answer that briefly? I think, Michelle, that you wanted—



HOUSE OF COMMONS

Dr Meadows: I can. We consulted on the principles of the algorithm; they were subject to public consultation in April. At the same time, we had set up a working group with technical representatives from the exam boards. The decisions taken were ours, but we worked with around 20 different technical experts from across the exam boards and then chose the algorithm that seemed to produce the most accurate results.

Q991 **Ian Mearns:** At what point did the algorithm mutate?

Dr Meadows: I don't believe that the algorithm ever mutated.

Q992 **Tom Hunt:** I am very interested in the comments you made about the possibility of socially distanced exams and how, early on, that was something you were thinking about. It seems like Germany obviously did progress with exams, and from what I have seen they did it pretty well; there do not seem to have been any major issues. Is Ofqual trying to look into what happened in Germany, to learn lessons to try to make sure that, if covid is with us for a while longer, we can try to eliminate any uncertainty about next year's exams as soon as possible? Finally, I know this is a bit of a direct question, but do you think, on balance, that it is quite regrettable that we could not progress with socially distanced exams and could not avoid these issues?

Chair: I will get the chair, Roger, to answer that, please.

Roger Taylor: Certainly we are conducting work right now to learn from international examples, and particularly the notion that there has to be some mechanism by which the individual candidate affects their fate by doing some kind of test of their skills and knowledge seems to us a key learning. Another key learning seems to be that, if you cannot really replicate your normal exam grades, do not pretend you can—a teacher certificate and other forms of representing information—because the objective here was to enable progress, not necessarily to award GCSE grades. What we needed to do was enable kids to move into university places. That is another key learning.

The exam question is precisely how it sounds. I think it is important to recognise that that cannot be considered independently of the degree to which schools are open. At the time we were looking at this originally, it was not clear how long schools were going to be closed for. In a situation in which schools are closed for a long time, it does make it difficult.

Chair: Thank you. Jonathan, can you do this quickly, because we have not had Christian in yet?

Q993 **Jonathan Gullis:** Yes. First, I want to quickly go back to the Chair's questions about the testing of this algorithm. Did you run the algorithm on the data that you had before the results were released, and therefore did you properly prepare the process for kids who want to appeal? How did that change after you received direction from the Secretary of State?

Secondly, my main concern is that you had teachers ranking kids within grades, and it appears that, the larger the cohort, the more that was kind of ignored. My partner, a head of RE, spent eight hours doing that with



her Department one day. Why was that ignored for large cohorts?

Roger Taylor: First, it was essential to have all that information. The actual grading provided by teachers was essential for us to be able to test different types of algorithm and see how they would work. They formed an important part of the quality assurance process. They were also an important piece of information in determining, for example, exactly what you are talking about—identifying this question of outlying candidates where it would appear that the change in grade produced by the algorithm was highly likely a reduction that could not be justified and would need to be overturned on appeal. The information absolutely was being used, but there has been an ongoing discussion about the level of weight that should be placed on the centre assessed grades.

The principles that we consulted on emphasised precision—accuracy of prediction—as being the cornerstone of what we wanted, and it was very clear that the variation between centres in their level of optimism did not reflect real differences in the achievements but other factors. We were conducting research with schools to understand how they would put their grades together, and we were clear that it would not have improved accuracy to have given greater weight to the centre assessed grades.

Perhaps one final thing to say about this, because some people have suggested that, if we had done that, the results would have been more acceptable: it is important to also recognise that that was the sort of approach taken in Scotland, and it did not make the results any more acceptable to individuals. On your specific point about the level of testing done, perhaps I can hand over to Michelle again.

Q994 **Chair:** Briefly. From what you said before, you did not do your own mock to see what would happen and then spend time rooting out the anomalies.

Dr Meadows: We did indeed know what the difference would be between the centre assessment grades and those produced by the model. We did indeed use a variety of approaches to look for where there were anomalous results. That is how we knew that in total it was about 0.2% of results that were potentially anomalous, hence we then thought very hard about how best to design the appeals mechanism so that we could deal with those anomalies.

Q995 **Chair:** But the appeals system, until the Government changed it, was still incredibly narrow. It was just based on whether there could be proof of bias or discrimination—something that we raised our concerns about when you appeared before our Committee.

Dr Meadows: The appeals system did allow a centre to come forward and say, “This year I have a student who would be expected to outperform compared to what we see in normal years. The shape of our grades this year—if you like, the grade distribution—will be different, and here is some evidence.” The appeals system did allow for that.

Q996 **Chair:** But you had said to us in your documents that it was based on bias



or discrimination.

Roger Taylor: Just to clarify the point, the individual student can appeal directly on the basis of bias or discrimination. If they believe the moderation process has resulted in their getting the wrong grade, they can ask the centre to make that appeal, but, crucially, the centre makes the appeal because often what was going to happen was that it was more likely the centre would be the organisation that recognised what happened and would put in the appeal immediately. It did not rely on the individual.

Chair: Okay. Christian, you have been waiting patiently. Go ahead.

Q997 **Christian Wakeford:** Thank you, Chair. We produced our report in July, over a month before the A-level results came out. As part of that report, we asked for the statistical model to be published so that it could be scrutinised. We asked for that to be done in good time. However, it was not published until the actual day of the results, so that did not add to the transparency of the issue. If anything, it heightened the anxiety of the students out there, especially when we looked at some of the criteria. Some subjects actually showed less than 50% accuracy in terms of marks. At what point did this flag concerns with yourselves that there were several subject areas that were anomalous in their own right because of the low level of accuracy?

Roger Taylor: Michelle, I will hand over to you in a minute. The primary constraints on transparency were, first, to not publish information that would influence the way teachers made their assessments. As soon as the centre test grades had been delivered, we published extensive information about how the algorithm would work. The limitation of what we published at that point was that we did not then wish to publish information that would allow individuals to identify the grades that they had been awarded, because the fair process is that all the students learn their grades on the same day. The final full information could not have been published until the day of the results without revealing to some students the grades they had achieved. Those are the fundamental limitations within which we were working. Michelle, do you want to add any further comment?

Dr Meadows: Yes; absolutely. The information that was not published until results day was essentially the algebra and the definition of what would be a small centre, which we did choose to hold back for the reasons that Roger has just described. In terms of accuracy, it did vary across the subjects. There is a benchmark that is used in assessment evidence that any assessment should be accurate for 90% of students plus or minus one grade. That is a standard benchmark. On average, the subjects were doing much better than that. For A-level we were looking at 98%; for GCSE we were looking at 96%, so we did take some solace from that. Our aim throughout has been to be transparent about the limitations of the model. That is exactly why we published those metrics on results day so that people could understand the limitations of what is possible with any kind of statistical moderation process such as this.

Q998 **Christian Wakeford:** Thank you. I have a few further questions, Chair.



HOUSE OF COMMONS

Although there is obviously a concern about grade inflation, what we saw was a massive scale grade deflation by the algorithm.

Roger, one of your own comments earlier in this session was that using the statistical and teacher rankings improved the level of fairness. Speaking to a number of students in my own constituency, that clearly wasn't the case; they would find that statement very hard to believe, and I feel much the same. I think of one local college where I think the accuracy was predicted at 40% being downgraded. I had one college with over 70% of their students being downgraded. That, to me, showed that there is a clear anomaly in that particular cohort, but then we have all heard horror stories of straight A students coming out with Ds or, in some cases, even Us. To get a U for an exam you have not sat because there were no exams is incomprehensible. I have actually seen these kinds of assessments from the school and their general reports; they were getting better results at GCSE than they got in their standardised marks for their A-level.

At what point did any of this get flagged up, that actually there is a serious, serious issue, and—to use their language—"I have got someone else's D," because of rankings, when quite clearly, looking at it, they were a straight A student for pretty much their entire academic career? This wasn't just the odd anomaly, or 0.2%, as Michelle raised. This is large cohorts of numbers that didn't slip through the cracks—the cracks gobbled them up in large numbers. It clearly wasn't fit for purpose. If there was more strenuous testing, as Jonathan highlighted, I don't think we would be in this mess now. So there is a clear failing at some point, and I just wonder what you think that actual failing was.

Chair: What you seem to be illustrating is a sort of charge of the Light Brigade mentality, because you knew that there were anomalies, according to what you have said to us, and in your statement, and yet you carried on regardless, because that is what the computer said you had to do. I think Christian, in a brilliant way, has summed up the general unjustness and unfairness of the system that you developed.

Roger Taylor: I think there are two different points being made, one of which I agree with and one I strongly disagree with. The notion that the algorithm was deficient and that somehow a better algorithm would have made this fair is exactly the wrong conclusion to draw from this, I think. The evidence that we have presented shows why—and there are two different ideas of fairness here. When I said it improved fairness, what I mean is it improved, for example, the likelihood that specific candidates getting their grades would not depend on the centre that they had been in, and that the differences between different socioeconomic groups were ironed out through the standardisation process, so in that sense it was fair.

It is important to say, when you say there were large numbers of anomalies, 98% of grades that were awarded were either the teacher-assessed grade, which was the majority of them, or were within one grade. So again, it is not true to say that there were very large numbers of large movements in grades that were unjustifiable. There were a very



HOUSE OF COMMONS

small number, which we felt could be addressed through the appeals process.

I disagree with the notion that this algorithm was not fit for purpose or that a better algorithm would have produced a different result; but I strongly agree with your statement that to say this was fair just fails to recognise what happens to students—just the level of accuracy that was fundamentally possible with the information that was available was too low to be acceptable to individuals, and we recognised this right at the outset. We identified this as a risk.

We knew that there would be candidates who would get a grade that would be probably one down from the grade they expected, that it would mean they would not have a university place, and that this would strike them as fundamentally unjust. There was the hope that the autumn series would be the response that would give them the chance to prove what they could do. That opportunity receded, which made the acceptance of the grades much harder; but your suggestion that to say that this was a fair way to award grades to students—it would only have been true if people had been prepared to accept it as the unfortunate consequences of a pandemic in a situation in which there was no other alternative. I think, quite reasonably, people were not prepared to accept that. That is precisely why I did apologise.

Chair: Ok. Christian, do you want to come back?

Q999 Christian Wakeford: I have one final question. I will agree to disagree with you on the algorithm, Roger; I am more than happy to invite you to my local sixth form college to let them explain it much more eloquently than I. My final point regards private students, especially those who are adult learners or in elective home education. Again, we referred to those in our report in July, and it seems that this is a large cohort that has essentially been ignored. The only avenue they have, because there is no centre assessed grade because there is no centre and no mock exams, is to do the autumn series of exams, in which case they have more than likely had very little input from a tutor or whatnot. Essentially, we are asking them to put their life on hold and be stuck in limbo for an exam that they will sit in the autumn, to maybe, if they are lucky, start a university course in January—although realistically, with larger numbers going this year, it is likely to be next year. Through no fault of their own, they are in limbo for a year with no real avenue to go down.

Roger Taylor: I have huge sympathy with these people. Clearly, they have been some of the people who have lost out most as a result of the decision to cancel exams. I will hand over to Julie to say a little bit more about this, but once the decision had been taken to cancel exams, it was very hard to find a solution. We explored extensive solutions, but ultimately the situation was one in which, once exams had been cancelled, these people had lost the opportunity to demonstrate their skills and knowledge in a way that would enable them to move forward with their lives. That was the situation we were in. I agree with you that it is a very



HOUSE OF COMMONS

tough situation for those people. Julie, do you want to say a little bit more about what we were able to do to help?

Julie Swan: Thank you. Again, one of the great worries of the model for us was how to bring in those students who did not have a relationship with a centre. In the absence of exams, we needed evidence from an exam centre about the likely performance of a student in those exams had they been able to take place. For students who had no relationship with a centre and no work with which a centre was familiar, that was problematic. The exam boards, though, did identify a couple of additional centres that were well-versed in working with adult and distance learners. The exam boards extended the deadline for the receipt of centre assessment grades to enable private candidates, where they did have some work they could draw on, to engage with that centre so that centre could get to know them, set some additional work for them and submit a centre assessment grade for them.

It is very difficult to know the precise figures, because private candidates come in many different shapes and forms and have different experiences behind them, but the best data we have suggests that about 3,300 private candidates did get an A-level grade this year. That is fewer than in a normal year, and it seems that the group of students who most missed out are perhaps students who would in another year have taken an A-level in a community language, perhaps the language they speak at home. They are not taught that language at school and therefore there would be no work on which their school or college could have drawn.

It has been a big worry for us throughout, and it was another reason why we were particularly concerned about the mock appeal route, because that could have been a further disadvantage for private candidates who might not have had recourse to that.

Q1000 **Kim Johnson:** I have three questions, all directed to Michelle. The Committee received communication late last night from Cambridge Assessment, in which they inform us that there were 12 different approaches considered by Ofqual, and that they offered some assistance but were not asked to review or scrutinise the final form. How true is it that Gavin Williamson received information about the flaws in grading two weeks before the results and that you gave assurance that this could be managed?

Dr Meadows: Thank you for the question. Right from the beginning when we put together a technical group to work on the model, we had all the exam boards but additionally technical colleagues from Cambridge Assessment. So colleagues from Cambridge Assessment have been involved throughout; they have had the opportunity to comment on the model and, I have to say, they have actually been very helpful in commenting on the model. They also, through OCR, had the opportunity again to raise concerns, in particular at the point at which the designed model was signed off. We have an approach where we bring together responsible officers from the exam boards who sign off the final



HOUSE OF COMMONS

requirements, and that was an opportunity for colleagues to raise concerns; none were raised. I want to be clear about that.

In terms of warnings that were discussed two weeks before, this was around anomalous results—outliers. Cambridge Assessment had used some of the techniques we created and some of their own to identify potential outlying candidates. This conversation was already going on between ourselves and Ministers, and you might remember that in the public domain there was debate about students who were, if you like, outliers, so it was very much a live issue. So there was no sense in which they had brought something to our attention that we had ignored; this was an ongoing conversation.

Chair: I think some people would disagree with you, like the many who expressed reservations to our Committee about the algorithm. We know that Sir John Coles was extremely worried about it—I will come to him in a bit—and we now know about the Cambridge Assessment group.

Q1001 **Kim Johnson:** We also mentioned flaws and concerns in our report in July. According to the Prime Minister, the fiasco was caused due to a mutant algorithm, as we discussed earlier. However, we have mentioned that your standardisation model did disadvantage certain groups, including high-achieving pupils in centres with lower historical attainment, black students and students in centres with larger cohorts. How did your testing process fail to identify the potential unfairness that appears to be inherent in the model?

Dr Meadows: The evidence that we published looking at the attainment gaps for, for example, black versus white, shows that there is no bias inherent in the model. I think there is a misunderstanding about what happened with the calculated grades, because if you compare the difference in outcomes for black versus white students, or indeed students receiving free school meals versus those who do not, with those for previous years, you will see that there is very, very little difference at all in outcomes for those groups. There simply is not evidence in the data of that kind of bias, but I understand that that has been widely misunderstood and misreported.

Kim Johnson: We received a lot of witness testimony from a number of our meetings, where people provided evidence about those disparities. What information has Ofqual collected on the effect of the standardisation process by area of the country? What comparisons do you have for the process between affluent and the most deprived wards?

Dr Meadows: You are absolutely right that there were great concerns about the potential for bias in this model, and it was a thing that we continually kept under review. You will remember that we put out additional guidance to teachers to support them in making objective judgments. We always worried that this was something we would have to address—that we would see something in the data that would suggest bias, and then we would have to revisit our previous decision that we would not set differential standards for different groups of students—but



HOUSE OF COMMONS

actually the data simply did not show it, so we did not need to review that decision, I am pleased to say.

We did publish information about calculated grade by region. That is all in the report that we published on A-level results day. We want to make a commitment to publishing further information and data on what the final grades this summer look like, so people can see that by region, and, as Roger has outlined, allowing everybody access—in a safe way—to all the data that we have available, so that people can reassure themselves that there has not been some sort of bias in process.

Q1002 Kim Johnson: Thanks, Michelle. How were the errors, including pupils on foundation GCSEs receiving impossibly high standardised grades, missed by Ofqual?

Dr Meadows: We took an early decision with regard to tiering. Tiering is, as you will know, for certain subjects where the span of ability of students is really wide and there is different content to be taught for the most able versus the less able. We have these tiers.

Normally what happens is that, if a student is entered, for example, into the higher tier and fails to achieve a grade 3, they fall off and become ungraded. For the foundation tier, there are a small number of students each year who are capped, effectively, so the most they can get is 5. That is just a product of the fact that they are put into separate papers.

In the absence of papers this year, we felt that the fairest thing to do was to remove those limits on students' performance. So there were a very small number of cases where, for the tiered qualifications, less than 1% of foundation tier students received higher grades and, for the higher tier, less than 0.5% received lower grades than they would normally achieve. We felt that it was a decision in favour of students—that they would not be constrained in the normal way.

Q1003 Chair: If you cannot get more than a grade 5, as I understand it, with the foundation GCSEs, how on earth can you develop a system that allows students to get higher than that? If you can answer that simply, so people can understand it, that would help.

Dr Meadows: We simply removed the cap that would normally be there, because that cap is an artificial cap. Unfortunately, it is a difficult decision for teachers whether to enter a borderline student for foundation or higher tier and sometimes, in a normal year, unfortunately it goes wrong and students do better than expected or worse than expected. We wanted to remove that constraint this year, so students were able to get higher grades through the standardisation process, and it did happen in a very small number of cases.

Q1004 Kim Johnson: Final question. I wanted to know whether an equality impact assessment was undertaken in terms of the process that you went down. If so, would it be possible to publish it? Thank you.

Dr Meadows: Yes, an equality impact assessment was done throughout for our various decisions. They have been published. We published one in



HOUSE OF COMMONS

association with the entirety of the process—collecting centre assessment grades and the standardisation process. Then of course there has been a great deal of equalities work as part of developing the model which, again, is published in our technical report.

Kim Johnson: Thanks, Michelle. They are all my questions.

Chair: Tom, you wanted to ask about BTECs.

Q1005 **Tom Hunt:** Michelle, students had to wait days for their BTEC results at short notice. Do you agree with the decision to award those qualifications in line with CAGs, as happened with A levels and GCSEs, or do you think the standardised grades would have been fair?

Dr Meadows: I think it is fair to say that public confidence and rebuilding public confidence in grades this year has been absolutely paramount, so we can understand Pearson's decision to want to revisit those grades, because what we would not want is a sense of unfairness developing between those students who took BTEC qualifications and those students who took A-level qualifications. In light of building that confidence, not wanting that sense of unfairness, I think we were happy that Pearson took this decision. It was a decision for Pearson, rather than for us; it was something that they wanted to do.

Q1006 **Chair:** Well, they had to do it, de facto, because of what had happened across the board with other exams, to ensure a level playing field.

Dr Meadows: I can see why that was a sensible thing to do, yes.

Q1007 **Jonathan Gullis:** In places like Stoke-on-Trent, we have a high number of BTEC students, and I need to get across to you the anger and disappointment. When I rocked up to schools on GCSE results day and was speaking to A-level students, they were having to have special things put next to their BTEC grade such as "To be confirmed" or slips inserted. That just makes a mockery of things. Again, I appreciate that Pearson have questions to answer, but I need to know from Michelle or Roger really, how much consultation was done with Pearson regarding the use of the algorithm, and subsequently, following the U-turn, how much information and one-on-one conversations did you have with them in the build-up to their decision? It seems mad to me that they made the decision to pull the results at 5 pm on the day before the GCSE results were announced, which caused huge upset and anger.

Chair: And, if I can speak to the chair, on the day that you made your U-turn announcement, the BTECs were almost like an afterthought—there was just a little line about BTECs. My impression was of a classic example of the way in which we regard vocational qualifications in our country—they are forgotten about. Even though 450,000 students were affected, because you were all so concentrating on the GCSEs and A-levels, you took your eye off the ball when it came to dealing with the BTEC problem, which happened inevitably once you had done the U-turn. Do you want to answer both my points and Jonathan's, chair?



HOUSE OF COMMONS

Roger Taylor: First, I want to say that I fully agree with the view that it was not acceptable for BTEC students to be put in that situation. I fully recognise that. I also want to clarify that there is a huge difference between the way that vocational qualifications work and the way that general qualifications work—

Q1008 **Chair:** No one is doubting that. Everyone knows that there is a massive difference. The issue was that, inevitably, there would have to be a change because of what had happened—because of the domino effect, in essence—but the BTECs seemed to be an afterthought, or hardly mentioned at all until it was realised that actually 450,000 students were potentially affected.

Roger Taylor: There are two points that I would take issue with. It was not inevitable that there would be a domino effect, because the use of calculated grades inside the BTEC system was completely different from what had gone on with general qualifications. They were two completely separate pieces: one Ofqual was closely involved with and where we had the authority to make a decision; and the second was one that Pearson were responsible for and where we had no authority to determine how they were going to respond to the situation. That was their call.

Those are the only points that I wanted to explain, in terms of how that operates, but I fully acknowledge that the consequence of that was that Pearson did decide to change their course. We understand why they did it, but the consequence was late grades for those students. I was pleased, obviously, that we managed to get the GCSE and revised A-level results out on the awarding day in that week, but that added to the pain for BTEC students, who were not able to have their results on the day.

Q1009 **Chair:** Don't you think Ofqual should apologise to all the BTEC students for what happened, because they had to wait longer than anyone else? Am I right that most of them got their results last Friday?

Jonathan Gullis: Even some of those were delayed, Chair. Results day was promised for the 25th, but then some results were after the 28th. It sounds like what you said earlier about the statement we got last night, Chair—"It ain't us, guv." That is what it feels like right now.

Roger Taylor: I am very happy to apologise for the consequences for BTEC students of having to wait.

To be clear—again, Julie might be able to provide more detail—my understanding was that the outstanding results were due to the information that Pearson needed to receive to be able to issue the results. Julie can you give us any more detail on where that is?

Julie Swan: A few results were issued later because the information needed to calculate the results was within the centre, and that had to be extracted from the centre. Because of the way in which BTECs are restricted and the availability of some assessment evidence, it was more complicated for the awarding body to re-run the result than it was for GCSEs. For GCSEs it was pretty straightforward—they had to replace the



HOUSE OF COMMONS

calculated rate with the CAG—but there was much more work involved in revising the grades for BTECs, hence the unfortunate delay.

Chair: Thank you. Before I bring in Apsana, who has waited a long time, Ian and Caroline, can you ask your questions briefly?

Q1010 Ian Mearns: This is in response to Kim's earlier question about the bias in the modelling. You said that the algorithm was not biased, but surely the context that led to the algorithm had bias within it, with the dominance of the history of results and outcomes that are biased. I understand that key stage 2 results might even have been taken into account. I am the chair of governors of a school and many of their pupils have English as a second language. Many of them have come into a primary school late in their primary-school life, but then flourish as their English develops in the secondary sector. That inherently has a hidden bias against youngsters for whom English is a second language. I wonder whether the algorithm could have been tested by running all of the 2019 results through it, just to see what would have happened.

Chair: Perhaps Michelle or Julie could answer that question as concisely as possible.

Dr Meadows: We did indeed run the 2019 results through the algorithm—that is how we tested its accuracy and the impact of the standardisation process on centres with varying proportions of students on free school meals, of varying ethnicities, and with English as an additional language, which was another thing that we looked at. We have looked specifically at English as an additional language and, again, there is not evidence for those students that the process itself produced bias. We have to remember of course that all those different groups perform quite differently in exams in a normal year because of the differences in educational opportunities. What we have seen is that those differences look like they do improve each year, so the process has not exacerbated them, but unfortunately, yes, they still exist.

Q1011 Ian Mearns: I just think that in future you should guard against using key stage 2 results at all because, as we have already heard from the Minister, Nick Gibb, in evidence earlier this year, key stage 2 results are not at all important.

Chair: Okay, I think that is an observation, but does anyone want to comment on that?

Roger Taylor: I think the one comment to make about that is that lots of issues have got wrapped together in this whole process, and that the fundamental objective here was to enable students to progress. One of the issues in that progression is the impact of different educational opportunities on different communities. Ofqual was attempting to replicate the most likely set of results that would have been achieved, which of course is replicating the consequences of differences in educational opportunity. It would not have been appropriate for us to have attempted to correct that. However, on teacher certificates or just thinking about if we had been addressing the problem of progression rather than the



HOUSE OF COMMONS

problem of awarding grades, it is important to say that we might have been able to think about this in a different way, which would perhaps have allowed us to think about the fairness of progression given the difference in educational opportunities available to young people.

Q1012 Apsana Begum: My questions are primarily for Julie. Your guidance on appeals does not allow centres who issued CAGs in line with previous performance to revisit those grades. What is your message to students in those centres who may have received lower grades relative to their peers?

Also, what are your views on what David Blow, the external advisory group member for Ofqual, has said about calling on the Government to allow for centres to resubmit these grades without reference to historical performance?

Julie Swan: We recognise the difficulties that some schools and colleges feel that they are now in. We have heard from some schools and colleges that they were very cautious when they submitted their centre assessment grades that they took into account the prior performance of students and took into account the prior attainment of this year's cohort, and that they were perhaps more conservative, if you like, and less optimistic about the grades they submitted than some neighbouring schools. And I think the concerns of those schools are that other schools and colleges were perhaps much more optimistic in the approach that they took, and that students in other centres have therefore got an advantage that has not been given to their students.

However, it is of course a moot point whether it actually is an advantage to be given a grade that is an inflated grade, which does not actually reflect the grades that you might have received if you had been able to take the exams. For example, we are hearing concerns coming from some further education colleges and some sixth-form colleges that students in centres that were much more optimistic about the grades that their students would have received have now perhaps—now they have got the centre assessment grades—got grades that indicate a level of ability that is not actually very accurate, and there are concerns that they might progress on to a course for which they are now ill-equipped.

So, as with so many aspects of this situation, there are pros and cons, advantages and disadvantages, to the approach that has been taken. Therefore, I think it is wrong to assume that students in the centres that were perhaps more realistic about the grades that their students would have achieved have necessarily been disadvantaged.

Q1013 Apsana Begum: I think it was mentioned earlier, but just for clarification for those who are watching, I think it was said that there was no requirement for schools to pre-moderate CAGs. Will that be guaranteed at the point of appeal?

Julie Swan: As we explained, teachers were asked to take into account a whole range of evidence in order to come to an holistic judgment about the grade that each individual student would most likely have achieved



HOUSE OF COMMONS

had they taken their exams. We did suggest a number of ways by which they could sense-check those views, including looking at the prior attainment of their students this year relative to the prior attainment of students in previous years, and the grades that their school or college normally receives.

I know that exam boards are having some discussions with some centres that believe they made a mistake when they submitted their centre assessment grades, and those are being put through the appeal process, where appropriate.

Q1014 Apsana Begum: Moving on, I want to bring up again our report in July, because we raised a number of serious concerns about the process for pupils who are unhappy with their CAGs and the general accessibility of the appeals process itself. Are you collecting data on which students are contesting their grades and how are you supporting students who have missed out because of their CAG?

Just for reference, there were a number of things that would be available for pupils and students that were mentioned in the evidence session that we had previously. So I just want to know how much of that has been implemented and whether data is going to be published on that as well.

Julie Swan: We were very aware of the concerns that the Committee raised in June and we published student-focused information about what to do if they had a concern about their grade, which we updated when the decision was made to rely on the centre assessment grades.

We set out some examples that might help a student to understand whether there is something funny about the grade that they have got, or whether, just as in any year, it is just a grade with which they are disappointed. We gave them some scenarios to illustrate the types of questions that they might be asking of themselves, their teachers and the exam board, to help them to understand whether there may have been evidence of bias or discrimination.

We published that information; as I say, we updated it when we took the change of position. We also briefed the National Careers Service, who run the exam results helpline and who were taking calls, as well as our own staff. We also provided information to the Equality Advisory and Support Service, who we know have also been taking calls from students who are concerned about their grades. We are having regular check-ins with the exam boards to understand the types of cases that are coming through as appeals presented by schools and colleges, or being raised by individual students as a concern or complaint about bias, discrimination or other forms of malpractice. We are talking to them on a very regular basis, and we will be collecting data.

Q1015 Apsana Begum: Can you just confirm again that the anonymised data on where appeals have been received will be published at the conclusion of the appeals process? We have mentioned in our report the breakdown of that data: for example, by type of school attended, geography,



HOUSE OF COMMONS

gender, ethnicity and FSM eligibility. Can you just confirm that that is going to be published, and when?

Julie Swan: We will be publishing data. I cannot yet confirm when, because I do not know when the process is going to conclude, and I cannot at this moment confirm the granularity of the data that we will publish. We will need to come back to you on that.

Q1016 **Apsana Begum:** One of the pieces of feedback I got from my constituents—we have a large number of students on free school meals, from really poor backgrounds—was that the National Careers Service was referring a lot of them to other numbers, other helplines to ring, and it was not really seen as a way in which they could get proper advice and guidance on appeals processes themselves. What would you say to that?

Julie Swan: I hadn't heard that feedback. We have been talking to the National Careers Service; we know they had a high volume of calls. We were getting a high volume of calls. We normally have two helplines open; we were running 10 helplines with extended hours around results day, to make sure we could give timely advice to students who had concerns. I will happily follow that up with the National Careers Service to find out more about those concerns.

Q1017 **Apsana Begum:** When we had Sally Collier give us evidence, she did say that she does not claim to have any easy answers to the issues around the appeals processes and accessibility. Again, she gave an assurance that students who feel they have been subject to bias or discrimination would be able to get support, but that they should approach their schools as well. Do you recognise the anxiety that many students and pupils feel about challenging a result that is determined by an institution? It puts them at unease, and navigating that process is very difficult.

We mentioned specifically that many pupils are unlikely to approach their MP, lobby, speak to lawyers and so forth. At that point, we were really unconvinced that the process for appealing against grades was going to be fair and accessible. How much more assurance can you give that the support that has been provided is accessible and reaching out to those who need it? Again, I do not have data to support this, but just looking at my own constituency, pupils have decided not to appeal because they do not think that the process is going to be worth going through, especially as people from ethnic minority backgrounds or socioeconomic backgrounds who just do not understand how to navigate through that process.

Julie Swan: That is why we spoke to the Equality Advisory and Support Service, and they were happy to have their information and their phone lines published in our guidance. We know that they have received a number of calls, as have the exam boards directly, and indeed Ofqual.

It is always tricky. I looked back at last year's data, and just for A-level results, there were 60,000 challenges to the marking of students' scripts in a normal year. In every year, some students are unfortunately disappointed with their grades, and the difficulty we have this year is



HOUSE OF COMMONS

unravelling whether a student is disappointed because something has gone wrong in the way in which their centre assessment grade was determined, or whether they are disappointed as they are, unfortunately, in the normal course of events. The scenarios we included in our guidance were intended to help students ask those questions of themselves and of their teachers, to help them work through whether something may have gone wrong for them or whether it was a case as in any other year.

Chair: Thank you.

Apsana Begum: Chair, can I just ask one more question?

Chair: Very briefly, because we are running out of time.

Q1018 **Apsana Begum:** On students with a SEND background, do you have any comments on how they are being supported through appeals processes, and particularly on the appropriate or inappropriate evidence submitted to determine grades?

Julie Swan: Yes, a couple of examples. We know from the exam boards that some students who would have had a reasonable adjustment when they took their exams have raised concerns that they do not believe that that was taken into account by their school or college. Some schools and colleges confirmed that, so the exam boards are looking at that in appeals. In terms of making sure that our information was accessible, you may have seen that our website produced a number of videos in British Sign Language to help BSL users navigate their way through the information. We also produced some easy reader guides on the approach.

Chair: Thank you. Christian, do you have any questions on the next part, because we may move predominantly to the autumn resits?

Christian Wakeford: I am happy to move on to the autumn resits.

Chair: Apsana, you have finished, is that right?

Apsana Begum: Yes.

Q1019 **Chair:** It was reported, chair, that you threatened to resign because the Secretary of State did not sufficiently show confidence in Ofqual. Is that correct?

Roger Taylor: I certainly spoke to the Secretary of State, and I think that, if the Secretary of State did not have confidence in the regulator, I would definitely need to consider my position.

Q1020 **Chair:** So did you say to him directly that you would resign?

Roger Taylor: I certainly felt that, if the Secretary of State had no confidence in me, it would be appropriate for me to consider that, but we discussed it and resolved that issue.

Chair: Okay, thank you. Caroline?

Q1021 **Dr Johnson:** If I start with what has been said so far, essentially, once



HOUSE OF COMMONS

the decision was taken to cancel exams, since 75% of children would in normal circumstances have received lower grades than predicted, the task that you were given of developing an algorithm that both commanded broad public support and was fair, and felt fair to each individual, was nigh-on impossible. Students going into year 13 this year and due to sit their A-levels next summer are very worried. What are you doing to ensure that students have a fair system for their assessments next year that commands broad public support?

Roger Taylor: I will start off, Julie, and then perhaps hand over to you. I think we have been very clear that we think that some form of examination or standardised test, or something that gives the student an ability to demonstrate their skills and knowledge, will be essential for any awarding system that the students regard as fair. We have done some consultation, and have published the results of that consultation, but it is obviously a fast-moving environment, and the impact of the pandemic remains uncertain over the future, so it is something that we are keeping under constant review. Julie, do you want to say anything more specific about that?

Julie Swan: Yes, certainly. Of course, we are just the exams regulator, so there is a limited amount that we can do to support students with their catch-up. However, we have given a lot of thought, and have consulted—we had a phenomenal response, as almost 29,000 people responded to our consultation—on what we can do on assessments, which is the bit of the system within our control, to both free up teaching time, where possible, and to make sure that assessments can take place fairly and within public health safeguards.

Content for GCSEs, AS and A-levels is of course determined by Ministers, and Ministers, as I am sure you will know, have agreed some changes to content for a couple of GCSE subjects—history, ancient history and English literature. We have published information about changes to assessment arrangements in other subjects that will free up teaching time, such as making the assessment of spoken language in modern foreign languages much less formal. The formality of the assessment and the need to record it can be quite disruptive to teaching. We know that, in the normal course of events, it can be accommodated, but in a climate in which every lesson and every hour of teaching will matter to students who need to catch up, we thought that that was a concession worth making, as well as allowing, for example, GCSE science students to observe practical science, rather than to undertake it themselves. We know from consultation that many—

Q1022 **Chair:** Julie, can I ask you to be as concise as possible? These are very long answers.

Julie Swan: Sorry. Lots of changes to lots of subjects. I don't know whether you want to come to the timetable, which is another issue.

Q1023 **Dr Johnson:** You talked about catch-up. We know the Government has invested £1 billion in tuition for students, particularly focused on those who are disadvantaged, to get people caught up. The more time teachers



HOUSE OF COMMONS

have, as you said, is important. I appreciate that if we delay the exams, you will have the marks in more quickly, but we do have eight or nine months to prepare. You have, therefore, eight or nine months to prepare people for the training and to mark exam papers, and to put this into place more quickly. How much delay do you think you could put into next year's exams and still have the marks in time?

Julie Swan: If you say "marked in time"—if you mean marks to meet the published results days—probably one or two weeks, but it would introduce risks. Anything more than that would be challenging. One of our concerns about delay is the willingness of teachers after an extended period of teaching, and having perhaps lost their holidays this year, to mark throughout the summer holidays, which, if the exam timetable is shifted back, is what would happen. That is certainly one of the major risks that we can see and that we would have to ensure could be mitigated.

Q1024 **Dr Johnson:** People do get paid for marking. It is a reasonably well-paid occupation, often undertaken by retired teachers at home, is it not?

Julie Swan: Most markers are actually active teachers. There is an army of teachers recruited—about 50,000. I don't think most teachers would say they do it for the money; they do it to make their contribution to the system. They might be willing next year, but they cannot be compelled to do so, although there are certainly opportunities to talk about incentives or the way in which schools and colleges could be supported to support their teachers to contribute in that way.

Roger Taylor: If I may add to that, we are obviously very conscious of the difficulties, and we are just expressing very clearly what the issues are with delay. I want to be really clear that, absolutely, we raised it in our initial consultation, and we are very conscious of the enormous benefit that would come from delay. We recognise the value in trying to find a way of making this work.

Q1025 **Dr Johnson:** For the exams themselves, at the moment students are supposed to be a minimum of 1.25 metres apart anyway for exam regulations. Presumably children would have to be 2 metres apart if you are doing a covid-secure exam. What work is being done already to ensure that we can have larger centres or more centres, more places for students to sit exams and more invigilators being trained ready? Given that it is likely that some young people may still be having to isolate or may be unwell on the day of the exam, have you considered the prospect of having to produce two papers, so that those students who are isolated through track and trace at the time of the exam can sit another, equally hard exam a couple of weeks down the line?

Julie Swan: That is a really interesting question, because that was in fact our preferred model for this summer: to run a normal exam series, if that was possible, with a set of extra papers in the expectation that every student, if they were ill during a period, would be able to take at least one paper in each subject, from which a grade could be determined. The autumn series will, of course, be much smaller than the next summer series, but we have advice through the DFE from Public Health England on



HOUSE OF COMMONS

the safeguards that will need to be put in place to run the exams safely. We will learn from that, but it is something that we are actively considering. It is something that needs to be taken into account when the timing of the exam period is finalised. One approach might be to have a later but condensed exam period, but that puts pressure on exam accommodation. Those are very much the issues that we are working through and will want to work through with every party who has a part to play in delivering the exam series.

Q1026 Dr Johnson: Just one more question, if I may. Obviously, this is a time of uncertainty for these students. Uncertainty provokes anxiety. When do you expect to be able to tell the students in year 13 what arrangements will be in place for their exams next year, even if it is a small range of options and the circumstances in which the different options may be used?

Julie Swan: In terms of changes to be made to assessments, those have already been announced. The timetable is not a decision for Ofqual alone, but it needs to be taken with Wales and Northern Ireland. We need to have a common timetable. We absolutely recognise the need for some certainty. We are working with the DFE to get to conclusions within weeks, rather than months. Several parties need to be involved in that decision.

Q1027 Chair: Just to confirm, did you make the case strongly that exams will continue to take place in the traditional way?

Roger Taylor: Yes, we have argued that we should try to make traditional exams happen.

Q1028 Chair: And you made that case strongly to the Secretary of State.

Roger Taylor: Yes, and we have published guidance on the various implementations that could have—

Chair: Could have taken place—okay. That is very helpful. Jonathan Gullis and Ian have some questions.

Q1029 Jonathan Gullis: Caroline asked some superb questions. My big concern going forward is that you have in education a body of teachers and students that does not have confidence in Ofqual right now. It is well and good saying that exams will be sat, but what happens where you have students in local lockdowns, where, God forbid, schools must shut, while other schools do not? I would like to know what work Ofqual is doing with exam boards to ensure that exams will be fair. As we know, teachers teach the curriculum in different stages and different orders, so that will be a challenge.

The second challenge will be plan B. Is it centre assessed grades, the reintroduction of coursework, a mixture of centre assessed grades and exams, or do you go to multiple choice? These are all questions. The Chair of the Committee made a good comment in the national press that these decisions must be made in the autumn half-term—the first half-term. I would be kind enough to extend to the Christmas break, but if we do not enter 2021 with a very clear set of directions for the teachers, the



HOUSE OF COMMONS

exam officers—who are some of the unsung heroes in all this mess, to be frank—and the students, then Ofqual must have plan B ready to go, which is likely to be centre assessed grades. As you can imagine, lots of students who are now in year 11 and year 13 are calling for the centre assessed grades to be used because they lost six months of education.

Chair: Just to come in on the back of that briefly, we should have thanked all the exam officers at the beginning of this session for all the work that they have done to support staff in very difficult circumstances. Also, God forbid, if exams cannot take place for one reason or another, would one option be to do the centre assessed grades, but to have an independent checker, perhaps a retired teacher or Ofsted inspectors, to check that they think those grades are fair? Is that feasible?

Roger Taylor: Just to say that I very much agree with Jonathan's comments about what needs to be achieved. I fully recognise and accept those comments.

In terms of a plan B based on centre assessed grades, our strong view would be not to do that. We do not think that there is a sensible mechanism whereby you can take highly variable evidence from a range of different circumstances and attempt to construct something that is a trustworthy way of discriminating between students based on their knowledge and skills. We think it is essential that students can take part in a fair comparative test that gives them the ability, on a level playing field, to demonstrate their skills and knowledge, and to be able to influence their own future.

Q1030 **Chair:** But if the exams cannot be done—what then?

Roger Taylor: We believe that there are mechanisms such as additional papers, and other mechanisms including using online tests. We feel we have enough time to come up with a solution to that problem.

Chair: Thank you. Ian, just before I bring you in, Tom Hunt had a question ready.

Q1031 **Tom Hunt:** Just to say I welcome your position on that. I think it is critical that we have exams next year and we should eliminate any discussion that there will not be. Plan A is exams must go ahead come what may. I think it is very good that you have that position.

My final question is on grade inflation. One reason why the initial approach was taken, in my understanding, is to guard against grade inflation. Are you concerned about the impact of grade inflation, both on those students who took the exams this year and on future years? What steps has Ofqual taken to try to guard against the dangers of grade inflation going forward?

Roger Taylor: We are very conscious of the effects. For this year's cohort, perhaps the main issue to consider is that universities may find they need to provide additional support to the candidates, and we may see some impact in terms of drop-out rates. We hope that does not happen, but we think that is something we all need to be alive to and think



HOUSE OF COMMONS

carefully about. The other obvious consequence of this is that, because many students are deferring their places, there will be an impact on next year's cohort, and we need to think very carefully about how we handle that in a way that is fair to next year's cohort. At this stage, I would say that we are very alive to those issues, and we need to build a system for 2021 that addresses exactly the issues you have raised.

Q1032 Ian Mearns: On that last issue of grade inflation, hasn't a decision been taken, though, to award students the centre assessed grade or the result after being run through the algorithm, whichever is higher? Doesn't that build in an element of grade inflation? Is that not the case?

Roger Taylor: That is correct. However, given that A-level grades have been awarded, we have a basic principle that, if you are awarded a grade, unless you have actually done something dishonest or cheated in some sort, you do not get downgraded as a result of subsequent events and appeals. That has been a fundamental principle of fairness, which was why the awarding of centre assessed grades was done in a way that did not result in anyone having a grade reduced.

Q1033 Ian Mearns: I have also had a significant number of people contact me, as I am sure have other members of the Committee, putting case studies, questions and individual problems to me, including one set of issues raised by the A-level grading issues support group—ALGI for short—and headteachers here in the north-east across several local authority areas. I am wondering whether I can forward all those questions and scenarios that have been put to me to you, so that you can look at them and give us some responses, please?

Roger Taylor: Please do. That would be enormously helpful for us to understand the issues that people are facing.

Ian Mearns: Lastly, as we have gone through today, the answers given have elicited a huge range of additional questions from the world outside via email, text and so on, so we may well be coming back to you and Ofqual, Roger, to get those questions answered.

Q1034 Chair: I have a few more questions. I wanted to raise the issue of those candidates, whether home educated or adult learners, who have no centre to predict or assess their grades. They have been left in no man's land. What are you going to be doing to look after those people—I think there are thousands of them—who are affected? What are you doing to help them?

Julie Swan: I said earlier that we recognise those students have been in a very difficult situation this year. The autumn exam series will be available for them, and that is one reason why we have said that exam boards must make exams available even if there is only one candidate wanting to take a paper in it.

Q1035 Chair: But we are talking about those people who are supposed to be predicted or assessed in terms of their grades and cannot be, because they do not have a centre assessed grade. There are thousands of people



HOUSE OF COMMONS

affected by that. What are you going to do to support those individuals and ensure that their grades are predicted or assessed?

Julie Swan: I am sorry—so, for students who had a centre assessment grade this year, they have had a grade issued. For private candidates who did not—

Q1036 **Chair:** I am talking about private candidates for exams who had no centre to predict or assess their grades and have been left in limbo. What are you doing to support them?

Julie Swan: The autumn exam series will be available to them. The DFE has written to centres setting out their expectations that centres will be welcoming of those students, to provide an opportunity for them to take exams within those centres.

Q1037 **Chair:** So they have to do exams in the autumn?

Julie Swan: Yes.

Q1038 **Chair:** So they are not going to be assessed in any way? Okay. We have talked about the many warnings you had about the algorithm. I have spoken to the very respected Sir Jon Coles on the telephone—he knows I am raising this with you—and he suggests that the 300-page document you produced has a large amount of system-level analysis, but very little analysis of the impact on students in particular types of centre or circumstance. His question was, in essence: what analysis was done on the impacts on high-achieving individuals in weaker centres?

Roger Taylor: We did a number of analyses. We constructed a number of different ways to think about this. We discussed it with Ministers and with the Department, and we shared analyses with the exam boards in order to understand precisely this issue. Michelle, do you want to say anything more about that?

Dr Meadows: That is absolutely right. We did indeed do these analyses. It is right that we did not publish them as part of our report, but I see no reason why we could not make that data available now and publish what we did so that people can see the approach that we took.

Q1039 **Chair:** So you did do the analysis of the impacts on the individual.

Dr Meadows: In particular looking for, as you describe, those students who had centre assessment grades that were, if you like, outlier from the other centre assessment grades within a school or college, and that therefore suggested that there was something unique about that individual and that the model might not work well for them. Exam boards would therefore approach the sixth-form college for a conversation to make sure that they were aware of the appeals opportunities. Yes, we did do that analysis, but it is right that we have not yet published it.

Q1040 **Chair:** Did you do any modelling of approaches that combined the centre assessed grades and the standardisation model? If you did, what did it show?



HOUSE OF COMMONS

Dr Meadows: From the outset in our consultation that we published in April, we discussed the balance that should be given in the model to centre assessment grades or the historical data. We were clear that, because of the different approaches taken by different schools and colleges in producing centre assessment grades, and the fact that some were far more generous than others, then to actually incorporate into the model and use both sources as evidence in an integrated way was probably going to be less fair. That was subject to public consultation. However, we did very much look at whether we could use the distribution centre assessment grades, the shape of that distribution, to inform the way in which the model could use calculated grades—

Q1041 **Chair:** So you did not do the modelling of the approaches combining CAGs and the standardisation model.

Dr Meadows: We did, but in a very particular form that I am concerned I am probably doing a bad job of describing. Essentially, if a centre was expecting a particular peak at grade A and then another peak at grade D, we could match that distribution of grades with the calculated grades, so we were looking at using both forms of evidence.

Q1042 **Chair:** If I could ask the chair, Roger, were Ministers made aware that any model of this sort, the algorithm, which kept the proportions achieving particular grades stable over time would inevitably mean large numbers of young people would be undergraded?

Roger Taylor: Yes. We said right at the outset that there was a certain level of accuracy that was achievable, although limited by the data available, and that there would inevitably be a group of candidates who would have done better in the exams, who would know in their own minds, and the teachers would feel completely confident that they would have done better in the exam, but that we would not be able to identify them, and—

Q1043 **Chair:** How was that made clear to Ministers? In letters? Meetings? Do you have the minutes of that?

Roger Taylor: It goes right back to the first paper we wrote.

Q1044 **Chair:** How many did you predict that this would happen to?

Roger Taylor: I will turn to Michelle, but given the levels of accuracy—I think this is right, Michelle; correct me if I am wrong—roughly a third of students would quite like to have a grade that was either one higher or one lower than they would have received had they taken the exam. It was simply not possible to determine which of them they were, so what we were doing was effectively finding a fair way of distributing that risk across the whole group of students.

Q1045 **Chair:** If you are going to answer, Michelle, can I just ask you additionally: the estimates of predictive accuracy—i.e. the 60% accuracy—assumed that schools' rank orderings would be perfectly correct. Would you agree that it is extremely unlikely that a school in a large-entry subject would get its rank ordering perfectly correct, and did



HOUSE OF COMMONS

you consider there was a way of mitigating this?

Dr Meadows: What we know from lots of research evidence is that teachers are much more able to accurately rank order than they are to make absolute judgments; but yes, you are absolutely right that there would inevitably have been some degree of error in their rank ordering. But actually, teachers are extremely good at it. It is impossible for us to know, of course, without some form of standardised testing, how inaccurate the rank orders would be.

On the issue of discussing changes to students' grades and the imperfections of any statistical model, I specifically remember during March discussing what we knew from A-level predictions for university entrance with Department officials and the Minister—that we know from that work that those A-level predictions do tend to be optimistic. A third tend to be optimistic, from the research evidence, and about a sixth are pessimistic. I think we knew from the outset that there would be large numbers of changes from what the teachers expected. There are several points at which there are papers, and so on, where we have set out the risks inherent in the model, whether that be—

Q1046 **Chair:** How much attention—either you, Michelle, or Roger—did you pay to the warnings of Sir Jon Coles and others who were very worried about what would happen to the algorithm, and school heads, and there is a whole list of organisations? What I am still not clear about is—I know you did a consultation, but I am talking specifically about the algorithm. When they warned about the algorithm and you didn't want to publish it in advance, even though we as a Committee suggested it, how seriously did you take the warnings of someone as credible as Sir Jon Coles? You can perhaps ignore the Select Committee, but Sir Jon Coles—a serious character; knows what he is doing.

Roger Taylor: First, I just want to stress that we certainly didn't ignore the Select Committee. We did publish details of the algorithm at the earliest opportunity we could. We took Jon Coles—

Q1047 **Chair:** You didn't, because we said you should publish the standardisation model well before, so that people could properly scrutinise it, and then you published it on results day.

Roger Taylor: Within the constraints that I set out around not enabling people to see their results in advance of results day—but, to take up the point about Jon Coles, we had a number of conversations with him. We took his comments very seriously. At the time we were doing a number of tests of different ways of incorporating centre assessed grades into the final award. This considered the issue, is the overall level of generosity something that we thought would reflect the real variation between centres? We felt that that didn't. We were then looking at whether the distribution of grades could be incorporated, but it was very clear that these were creating more inaccuracies than accuracies, because the distribution of grades was tending to be influenced far more by this knowledge of the—



Q1048 Chair: The reason for my question is that you had so many warnings from very good people without axes to grind, and yet, as I say, like the charge of the Light Brigade, you just went on and on and on, and as a result we came to the position that we did.

Roger Taylor: I disagree with that in two ways. First, I disagree that we just went on. We did not. We stopped and we considered these things very carefully. The second point, and perhaps the more important point, is I think there is some conclusion you are drawing, which is that if the model had been different and had been more in line with what Sir Jon Coles suggested, we would not be in this difficulty. I think that is simply not supported by the evidence. Particularly, I would, again, point to the fact that in Scotland they did adopt an approach that was much more one of combining centre assessed grades with the statistical predictions, and these did not produce any greater public acceptability of the results, because the fundamental reason that the results were unacceptable is for the reasons described earlier.

Q1049 Chair: I am very happy if Dame Glenys answers the next two questions. One of the complaints I had from members of the public, schools and particularly the press was that, when all this happened, basically Ofqual shut themselves down and refused to talk to the media. In fact, Kate Ferguson, a senior journalist on *The Sun* newspaper, and others, tweeted about it. Almost every education journalist I spoke to said it was impossible to get through and everything seemed to be referred to the Department. Why did that happen? What went wrong with your communication? Shouldn't you have been out there reassuring people rather than hiding away? Also, your former chief exec did not give any interviews or say anything publicly before she resigned from her position.

Roger Taylor: In the days just before and after the awarding of A-levels, as I have explained, we were in a difficult conversation with the Department for Education, which had announced a new policy with regard to awarding and were trying to resolve this. The announcement had been made to support public confidence. We were at variance about the best way to introduce this policy, and we felt very strongly that it would not be helpful for us to be talking to the media until this had been clearly resolved—

Q1050 Chair: But was that not unacceptable because everybody was anxious? You are an independent organisation; people read the newspapers, all of which were referred to the Department for Education. I am looking at your organogram—again, perhaps this is something that Dame Glenys may be happy to answer—and you have got one director of communications with 10 people working under that director, and according to Ross Kempell in *The Times*, you are about to hire another senior PR person for £80,000 a year. Given that you are spending all this money on communications—by the way, I do question whether all that is value for money to the taxpayer—why on earth did you hide yourselves away in the Ofqual attic, refuse to reassure anxious parents and the public, and ignore the media across the land, who were trying to explain what was going on?



HOUSE OF COMMONS

Dame Glenys Stacey: If I could just say, it is not for me to speak on behalf of the former chief regulator, but—*[Interruption.]*

Chair: I am sorry, Dame Glenys, but your sound is not working very well. Perhaps broadcast can fix that, if it is possible. Roger, do you want to answer?

Roger Taylor: I would like to give Dame Glenys a chance to answer this, but I do think it is really important that communication has to be able to effectively address people's anxieties and give them the reassurance they need. That is the fundamental principle that we were seeking to achieve through any—

Q1051 **Chair:** But you didn't. I have listened to you very fairly, but I think your refusal to engage with the media was genuinely shocking, given all the money you spend on communication officers. We will look at that as a Committee at a future date. I do not understand why you are hiring yet another one if you have got 11 of them—or why you need another one for 80 grand a year. I think it is pretty poor that you refused to engage with the public, the media and schools publicly during the controversy, because all it did was sow more confusion. You say you are independent, but everything was referred to the Department for Education. It questions your independence if you are only referring things to the DFE.

Roger Taylor: I do not agree with that. I think that if we had entered the debate at that point, it would have increased the level of confusion. Often, the best PR advice is not to speak until you can speak with authority and clarity. I think that is what we were doing.

Q1052 **Chair:** I think you should have answered questions to the press—to respected education editors and correspondents in the media. I genuinely cannot understand why you took that decision or why your chief executive did not go publicly and say sorry very early on when it looked like it was turning into a disaster. Dame Glenys, I do not know if your sound is working.

Dame Glenys Stacey: I hope you can hear me now. It was just to say that, like you, I have a strong interest in our communications capacity here and I do have that structure and arrangements under active review, just to assure you. Thank you.

Q1053 **Chair:** Eleven communications officers sounds a lot for an organisation such as yours.

Dame Glenys Stacey: Just to be clear, a lot of those are fairly junior ranks—forgive me—and they are predominantly involved in the production of significant standard reports.

Chair: Given that you refer everything to the Department for Education, you might as well just let the Department for Education handle your press issues. Why do you need another one on 80 grand, on top of all the others you have? I just question that.

Caroline, I will bring you in, and then I have one final question.



HOUSE OF COMMONS

Q1054 **Dr Johnson:** Going back to the different analysis that you did, my understanding is that there are a number of different exam boards and, for example, when you get a GCSE in maths, it could be an AQA GCSE, an Edexcel GCSE or an IGCSE, but what you get on the paperwork is essentially the same. I also understand that different types of institutions, different types of schools, are likely to use different exam boards. When you were doing your analysis, did you look at that and did it have any effect on the results?

Dr Meadows: This year, all the data was compiled together, so that one model for a subject was run across all exam boards. It would have made absolutely no difference to a centre which exam board they were with.

Q1055 **Chair:** Thank you. I have a final question. Given what has transpired—again, I am very happy if the new chief exec answers this—is Ofqual fit for purpose? Should it be brought back into the Department for Education? Will both the chair and the chief executive answer?

Roger Taylor: Dame Glenys, I will say a little first, then hand over to you. First, this year has obviously been a major blow to confidence in Ofqual. However, over the previous decade, since Ofqual was created, we have successfully brought in new GCSEs, we have transformed the regulation of vocational qualifications and we have started to use our regulatory authority to improve the security of examinations and their rigour. Prior to this summer, we demonstrated the real value of an independent regulatory organisation. Clearly, the organisation's reputation has taken an enormous knock from what has happened this summer. That is why we have brought in an extraordinary arrangement in order to help rebuild confidence in the organisation, including—I am very pleased to say—Dame Glenys's agreement to step in as acting chief regulator. Dame Glenys, perhaps you would like to say a little more.

Dame Glenys Stacey: What we have seen and heard in an excruciatingly difficult situation this year during the pandemic has shown the importance of an independent regulator. That is one of the reasons that I am here: I believe in it, and I know that there is all-party agreement with the notion that the work of the regulator should be truly independent of Government at the time of being set up.

Q1056 **Chair:** So you believe that the organisation is fit for purpose. Is any reform needed?

Dame Glenys Stacey: This is a small organisation. It is doing a very big job. We have spoken mostly about general qualifications today, but there are 15,000 or so vocational and technical qualifications, 150 awarding bodies that we oversee, and a very large-scale reform programme for vocational and technical as well. We have very important and very rare skills and expertise here. Yes, we have had an enormously difficult year, which I know that the organisation is trying to be as open and as sorry as possible about, but we have an important job to do and we need to get on and be able to do it.

Q1057 **Chair:** Will you make reforms?



HOUSE OF COMMONS

Dame Glenys Stacey: There are questions that we must ask ourselves about how we got to this position. You raised communications as one of the issues, and certainly we were heads down trying to do our most difficult, if not impossible, job. I think now that one of the lessons of this is that we need to have wider discussions as we face 2021—for example, in relation to the timing of results, how those affect the wider system, or higher education, and so on. There must certainly be a broadening of this engagement, in an adult way. We need to play our full part. Yes, I will look at the way in which Ofqual works—of course I will. If any changes are needed, I will put those in play, even as an acting chief regulator.

Q1058 **Chair:** Ensuring both quality and value for money for the taxpayer.

Dame Glenys Stacey: I have got that message loud and clear. Thank you.

Chair: Ian, you have a question.

Q1059 **Ian Mearns:** I am just wondering, therefore, whether one of the issues that you really should address, Dame Glenys, is that 25% of grades on an annual basis are regarded as being unreliable, either up or down.

Dame Glenys Stacey: Thank you. I'll certainly keep that in mind, and I look forward to speaking to you about it, Mr Mearns. It is interesting how much faith we put in examination and the grade that comes out of that. We know from research, as I think Michelle mentioned, that we have faith in them, but they are reliable to one grade either way. We have great expectations of assessment in this country.

Chair: Thank you. I appreciate your coming and at least being accountable for a very long session of the Committee. At the end of the day, you are public servants. I wish you all well and appreciate very much that you have been accountable, as I have just described.