



Education Committee

Oral evidence: [The work of Ofqual](#), HC 403

Wednesday 14 October 2015

Ordered by the House of Commons to be published on 14 October 2015.

Written evidence from witnesses:

- [Ofqual](#)

[Watch the meeting](#)

Members present: Neil Carmichael (Chair), Lucy Allan, Ian Austin, Michelle Donelan, Suella Fernandes, Lucy Frazer, Kate Hollern, Ian Mearns, Caroline Nokes, Kate Osamor.

Questions 1 – 69

Witnesses: **Glenys Stacey**, Chief Regulator and Chief Executive, Ofqual, and **Amanda Spielman**, Chair, Ofqual, gave evidence.

Q1 Chair: Good morning and welcome to our Committee. We have a morning basically to probe the work of Ofqual. It is great to see you both here and thank you very much indeed for coming along. We have virtually a full line-up, if not a full line-up, of members of this Committee, which is rather good.

I am going to start off with question No. 1. Basically, it is that Ofqual's perception survey and written evidence to the Committee suggests that the reform of GCSEs and A levels has been implemented a bit too quickly, with little consideration of the concerns raised throughout the consultation process. Is Ofqual doing enough to support teachers, students and employers to understand the impact of these reforms?

Glenys Stacey: Thank you very much and good morning. We are playing our full part in informing schools, colleges and students, parents and employers as to the nature of the changes. I have brought with me, for example, a recent publication, a set of what we call postcards, explaining the key aspects of the reforms. We have sent that with an accompanying letter to some 8,000 schools at the start of the term. It is part of an increased effort on our part to get the key information across, but of course we are not the only ones responsible for informing schools. Schools in turn rely very much on exam boards, quite properly so, and exam boards are doing significant work to make sure that schools understand the detail of the changes for each of the specifications coming into schools for first teaching this year.

Q2 Chair: Do you think this document was issued in a timely fashion, given the fact that it arrived in September?

Glenys Stacey: It is not the only thing that we have done, Chairman, so there are other efforts that we have made. I have written to schools before, and of course I appear regularly at events to speak, especially those events that teachers attend, and we work closely with HMC, ASCL and NAHT to make sure there is a good exchange of information about what the issues might be. An example of that is in recent months we have been doing good work with ASCL about teacher concerns and interests in the national reference test and making sure that our messaging on that deals fairly and squarely with concerns that are expressed. That is what we have been doing. It is right to say that this is the first publication of this nature, but it is certainly not the first thing that we have done.

Q3 Chair: Okay, thank you. In written evidence to this Committee, AQA suggested that Ofqual should allow exam boards to assist in training of teachers on new GCSE and A-level specifications. That was their clear recommendation. Will you allow exam boards to use experienced senior examiners to provide online training to schools and colleges, helping to guide teachers through the new qualification and specifications?

Glenys Stacey: You will recollect, Chairman, given your history on the Committee, the difficulty we had with face-to-face interactions between chief examiners and teachers, which culminated in what I think we call “Seminargate”, which was basically a specific problem identified with WJEC, where an examiner who was in GCSE IT gave advance notice of a live assessment question. You will know that the Committee at the time welcomed the steps we then took to strengthen controls and to effectively ban that immediate exchange because of the risks around it. It is true to say that the risks versus the benefits of direct exchanges, I think the balance has changed since that time. We have committed to reviewing that and will certainly be actively considering it now, looking ahead of Christmas, as to whether any changes need to take place.

Q4 Chair: Yes, because it would have to be done with reference to that discussion that we had.

Glenys Stacey: Absolutely, so it is just a re-evaluation of the risks and the cross-checks we can put in to make sure that we manage the risk of undue information being exchanged, but we absolutely want to support teachers in being as ready as possible, and we know that that requires close interaction with exam boards and the detail of those specifications. It is about how that is best done.

Q5 Chair: Striking the balance?

Glenys Stacey: Yes, and web-based technologies certainly have their place, as I am sure AQA would have argued.

Chair: Thank you very much. Ian, you are going to be talking about GCSE grading.

Q6 Ian Austin: I am, yes. I want to ask why this move from letters to numbers? I think the letter system is pretty well understood by parents and young people, by the

profession and employers. I think the issue is consistency year-on-year and between subjects and exams and there is some equality. That is what people want to know, so why the shift?

Glenys Stacey: All right, thank you. First of all, there have been numbers in the past, so these things do shift on occasion, so it is not the first time that we have had a shift or indeed used numbers.

Q7 Ian Austin: Sorry to interrupt. When did we last?

Glenys Stacey: Quite a time ago, so I have directives—

Amanda Spielman: In 1975.

Q8 Ian Austin: That is a long time ago, is it not?

Glenys Stacey: Yes. It does not happen every decade, but when there is a significant change in qualifications, and this is a very significant change, it is not uncommon to change the nomenclature. One of the reasons for that is so that in years to come, in 10 years' time when you are looking at someone's CV, you could see from the qualification grade that they were in the cadre of new GCSEs or old GCSEs. It signals the change and makes that change transparent and apparent. That is one reason for doing it, but not the sole reason.

The second reason is that we have moved from eight grades to nine, so we have greater differentiation across the grade spectrum in the new GCSEs and a slight shift towards more differentiation at the top end, which was what was being requested by the users of those qualifications. Therefore, if you move from letters to numbers you have a nine-grade scale with numbers. If you add another letter to a letter scale to get that greater differentiation then there is added confusion, because people are trying to reconcile old letters to new letters when they do not quite equate.

The last thing to say is we did consult on the change and we asked people in particular whether they wanted the numbers 1 to 9 or 9 to 1, and they went for the order we have, which is basically 9 at the top end. Those are the reasons why the changes were made.

Q9 Ian Austin: Okay. I am not persuaded, to be honest. I think it is going to be quite confusing, particularly during the introduction. The way it is being introduced, with some subjects coming in first and others later, means it is going to be quite a lengthy period when people are going to get a mixture of numbers and letters. How can we be confident that anyone is going to understand that—employers or institutions that students might be going on to? How can we be confident that employers, certainly while it is being introduced, will not discriminate against students who have numbers because they do not understand what the new system is when they do understand the old system?

Glenys Stacey: Okay, thank you. Certainly Ofqual, but also others, have a job now to do to make sure that there is a sufficient understanding of these changes in the grading structure. We know from the surveys that we do that there is more work for us to do there. We have a nice table, as you know, showing how each of the grades sits and equates and we will be pressing that home more. Certainly we have more work to do with those who represent employers in particular to make sure that change is understood.

Having said that, I think once it is explained it is relatively straightforward to see how the two things relate. I do not think it is impossible for most of us to grasp the change, whether or not one accepts the rationale for it. Seeing what old and new look like and how they equate and compare is not difficult, so we have to just keep pressing those messages home.

Q10 Ian Austin: How long do you think it will take before someone running a small business in the midlands understands that the system has changed from letters to numbers and knows what the numbers are? I think people have no clue that this is happening.

Amanda Spielman: A couple of things here. One is to amplify what Glenys just said, but the concerns you raise are exactly why we are making sure that in the transition there is a very clear read-across from the key A and C grades into the new system, so the bottom of a C will match the bottom of a 4 and the bottom of an A will match the bottom of a 7. When we are considering the different ways of doing it, that is exactly why—to provide that read-across.

The second thing is that fewer and fewer people go straight from GCSEs into the labour market. The proportion of people who transition without an FE course and a grade, an outcome, from a subsequent big course is tiny, so it is part of a bigger picture. This is not the sole basis on which these people are going into work.

Q11 Ian Austin: All right, but I guarantee there will be lots of children telling their parents that 1 is the highest and 9 is the lowest.

Can I just ask one other thing very quickly, which is on the current system? A head teacher has told me that it is possible at the moment to get a C with getting as few as 18% of questions right. Is that true?

Glenys Stacey: The current position is that—my staff will correct me if I get this wrong—generally speaking the average position is that a C equates to somewhere around 55% to 57% of marks.

Ian Austin: So it is not true that you could get a C on—

Glenys Stacey: No, that is the general position, the average position. There will occasionally be outliers, so perhaps in a low take-up subject, a rarely-taken subject, for example, and where the paper has not performed as well as it should. That might then produce low-grade boundaries or high-grade boundaries. Exam boards learn from that experience, do the post-award evaluation and make every effort to produce a better assessment next time. It is not that it cannot happen, but that it does not generally happen. In fact, what we normally see is quite a reasonable distribution of marks.

Q12 Ian Austin: Would it be possible for you to let us have information on how low and in what subjects it is possible to achieve—

Glenys Stacey: Yes, absolutely. Just to give you some assurance now, I am not aware of boundaries of the nature that you suggest in recent times, but obviously I would like to check it.

Ian Austin: Thank you. I have been told it is current.

Amanda Spielman: I can give you chapter and verse on the particular example that is widely quoted and requoted. That was from the days more than a decade ago when mathematics had three tiers at GCSE and the top tier was very definitely aimed at above-average ability students, but offered effectively a compensatory C grade for people who should not have taken the top tier. It was in that exam, which was not aimed at C grade students at all, that that 18% boundary happened, I think more like 15 years ago.

Glenys Stacey: That is right.

Amanda Spielman: It is understandable that the story circulates, because it is not an ideal place to be, but it is not representative of where we are now.

Glenys Stacey: It is one of the enduring stories within education, is it not? But, as I say, I am not aware of anything equivalent in recent years. I would be happy to check back and let you know.

Q13 Lucy Frazer: What do you think is a good pass in the 1 to 9?

Glenys Stacey: Just to answer that, we have said publicly that we will equate as clearly as possible grade 4 with grade C, or the bottom of these grades, and we have that clearly drawn here. It is important to remember that these grades are all pass grades, so the qualification is to be as accessible as it is now across a broad range of students.

Q14 Lucy Frazer: So a good pass in your thing is 5 and above, but I think Amanda said earlier that bottom C would be a level 4, bottom of 4, so a 5 is lower than a C?

Amanda Spielman: No, it is higher. A 5 is higher.

Glenys Stacey: So we are going up the grade.

Q15 Lucy Frazer: Oh, we are going up the grade?

Glenys Stacey: A good pass is effectively a performance measure. It is setting the expectation. It is the Department for Education setting the expectation for schools and how schools are to be measured. The important thing from our perspective of course is that we hold standards steady and make plain where we are effectively hardwiring the standard.

Amanda Spielman: The way to think about our job is that our job is to make sure that there is a good yardstick that does not get stretched or bent; it is not to say what is a good place to be on the yardstick. That is for Government.

Chair: Thank you. Suella, you are going to be asking about accreditation process.

Q16 Suella Fernandes: Yes. Could you just very briefly summarise the relevant criteria for accreditation and the process?

Glenys Stacey: Of course. Just by way of background, the Committee has always been an ardent supporter of accreditation and recognises it as a key control for GCSEs, AS and A-levels. We have, ahead of the reforms that are now going through, reviewed and strengthened our accreditation process, so we have now a cadre of some 650 subject experts, for example. Their details are published on our website and they play their full part in the process.

If I summarise it, there is a submission put in by exam boards for each of the specifications they wish to be accredited. That submission will have full details of the qualification, it will have some sample assessment materials—dummy papers, effectively—and the mark schemes that would be used for those papers. It also has—and this is a new development that we have driven this time—an assessment strategy for the qualification itself, which is setting out for us in detail how the exam board will assure itself and us of the validity of the qualification going forward. That deals with things like how they will make sure there is sufficient curriculum coverage in each of the assessments as they run, how they would avoid problematic predictability but assure sufficient predictability, so it is striking that balance, how they have approached the design of the questions and the balance of closed to open items, that sort of thing, and indeed their approach to the design of the mark scheme. So we get a good insight into the approach to ensure that this is a good qualification.

We then have subject experts looking at it ahead of time, and then they meet together with our staff as well, usually for several days in our offices or in local premises. They make recommendations to us. We then have a separate responsibility within Ofqual with a separate decision-maker deciding whether or not the specification and the accompanying documentation has satisfied us that that board can produce a good qualification and an assessment that will function sufficiently well in 2017. It is a resource-intensive process, but we do regard it as a critical control.

Q17 Suella Fernandes: Can you give me an idea of the figures? How many submissions have you had and how many boards are submitting?

Glenys Stacey: Again, I will be corrected if I am wrong, but my recollection is that this is another tranche of submissions. Of course we went through the whole process last year for the first round of reforms and we have started the first teaching now. We are on a second tranche now and my recollection is 156. Is it?

Amanda Spielman: Yes.

Glenys Stacey: Yes, 156 submissions have been submitted for accreditation, so we have gone through that process 156 times.

Q18 Suella Fernandes: Right, and how many boards are you generally regulating?

Glenys Stacey: Four. Now, not all exam boards are producing all the subjects and some will produce at GCSE but not AS or A level, so it is a mixed economy.

Q19 Suella Fernandes: There are more exam boards than four though, are there not?

Glenys Stacey: No.

Suella Fernandes: There are just four exam boards?

Glenys Stacey: Four of the awarding bodies that we regulate produce GCSE, AS and A-levels and are able to do so. We recognise them to do that. We regulate about 160 awarding bodies overall, but the vast majority are not in the GCSE, AS, and A-level market.

Q20 Suella Fernandes: Okay. In terms of the process and implementation there have been problems, have there not, with the implementation of the new GCSEs and A levels? For example, it appears that only Latin is currently the subject that is accredited and there is a backlog with accreditation of other subjects.

Glenys Stacey: As of today, there are three subjects accredited. That may not sound many, but it is more than one.

Q21 Suella Fernandes: Out of?

Glenys Stacey: Out of the 156. But when we look at where we were this time last year, we went through a similar process last year, when there was a very low rate of success on first submission. We plan for and expect several submissions. It is an incremental process, we find. Indeed, in days gone by, before we approached the regulation of such reforms in the way we do now, we would routinely get five or six submissions for any qualification. We plan for several submissions, so this is not unexpected.

What we do after the first submission is give examples, very detailed feedback, on every aspect of their assessment strategy, their sample materials, the individual questions in those materials, and indeed any aspect of the submission that we have a concern about. We know from speaking with exam boards that they value that feedback highly. They then rework their submission and resubmit. Looking at the resubmissions that we are now receiving, we are seeing considerable improvements, which is what we would expect. Given the nature of the process, it would encourage that, and that is what we are seeing.

Q22 Suella Fernandes: What is your view of the amount of resources that you have to deal with this task? Do you think it is sufficient based on the rate of accreditation that you are currently achieving?

Glenys Stacey: Thank you. So all regulators are in the same position, and we are no different. You have the resources you have. We have something like 160 staff, for example, and that is our main resource. We have to decide our priorities. So, yes, we regulate for reform, which includes accreditation, but we also regulate existing GCSEs, AS and A levels, and of course those other awarding bodies I mentioned, and all of their qualifications as well. We are always going to be striking a balance as to where we put our resource. It is absolutely right, in my view, that we give priority to accreditation. We are doing that, so it comes first at the moment. While we are in this intense phase it is quite demanding on the organisation, but that is the nature of things. We do not complain about that; it is the way things are.

Q23 Suella Fernandes: Just two more points I want to raise. The other problem that was highlighted by *The Times Education Supplement* was the abandonment of the extra checks on the science. That does serve to undermine confidence in the accreditation of that subject and overall. What is your response to that?

Glenys Stacey: My response is I thought it was a pretty irresponsible piece in *The Times Educational Supplement* and it was incorrect. Indeed, we have written to the editor to explain the truth of the matter.¹ The truth of the matter is that we did four strands of

¹ Glenys Stacey subsequently contacted the Committee to clarify that the TES article had been based on an error made by Ofqual. She has apologised to the Committee and the editor of TES for the mistake.

research to support maths accreditation. One of those strands was making comparative judgments about the expected level of difficulty of questions, using people expert in the subject. We chose to run a similar piece of research for science, recognising of course that the spread of questions and the type of questions that one has are much broader in nature than the sort of questions we get in maths. So we were aware that we would not get a good correlation. We got the results back from the research and, much as anticipated, it was not as strong a correlation as we saw in maths but it was not of no use. It was of some use to us, and indeed that information has been fed into the accreditation process for science and it has enabled us to think more carefully about the holistic nature of the judgments that we make in accreditation, and indeed the data is currently being further analysed so that we can get further information and assistance out of it.

I think what is important here for those listening is that this is the nature of research. We are at the leading edge in the research we are doing here ahead of these qualifications running and one is always looking for research to tell you what you need to know, using the best research techniques that one can. But you have to be realistic about what it is going to tell you and the key is in interpreting it intelligently. That is, I believe, what Ofqual does: we do good research and interpret it with intelligence.

Q24 Suella Fernandes: Last point. I think the consequence of those issues that I am having, those problems, is that there is lack of confidence in the coherence of exam qualifications and there is variability. If you do science with one board, you have a better chance of getting an A because it is effectively easier than doing science with another board, and you are getting a race to the bottom. Boards are competing to get more customers so they will have a higher success rate and that is only going to serve to undermine the legitimacy in the whole system. Do you not think there is an argument for increasing more rigour in standardisation and fewer exam boards?

Glenys Stacey: Right, so if I could just deal with the first points you make about confidence, because I get out to schools a fair amount—indeed, I was at a school just yesterday—and the impression I get is that schools, although there will be understandable frustrations about the speed of reform and so on, no one is questioning the accreditation process. People understand entirely the rigour we are applying to that and they respect what we are doing—that we are not letting qualifications through until we believe they are of the right standard and they are sufficiently comparable. We are moving into a fundamentally different way of regulating here to ensure sufficient comparability.

Secondly, it is right to say that there is a concern about a race to the bottom, and indeed this Committee has found before in earlier inquiries and said their legitimate concern is about a race to the bottom on subject standards, so is the subject sufficiently covered or are there easy routes through? I think those in the schools would recognise and agree with me when I say that we have closed off those avenues. In these new qualifications we are requiring sufficient of the curriculum to be covered in every assessment. That is what the assessment strategy requirement is about and why we are looking so closely at those sample assessment materials. We are making sure those assessments do the job, and indeed, we are doing unprecedented work in some subjects to make sure that they are sufficiently comparable.

If we look at the maths research that we have done and then the changes that we required from each of the exam boards to their maths specifications, there is no doubt, and the

evidence is compelling, that those specifications are at an unparalleled level of comparability in terms of expected difficulty. You cannot get a hair's breadth between them. That, as I say, is unprecedented. There is a fair amount of work and effort going on to make sure that these are as comparable as they can be.

The question of whether you have four boards or a different model is a matter for Government and of course that can be considered. Our job, of course, is to regulate the model we have.

Chair: Thank you very much. That debate has been raised by Nick Gibb and others of late and we will be thinking about that in our next inquiry: are qualifications and curriculum fit for purpose? Kate, you are going to talk about science in particular.

Q25 Kate Hollern: Absolutely. Your decision to remove the assessment of practical work from GCSEs was met with criticism both from DfE and scientists. Is it appropriate for Ofqual to make such controversial decisions in the face of such strong opposition?

Glenys Stacey: Thank you. I will start on that and perhaps you might like to come in. First of all, the Secretary of State was certainly very concerned that we spoke to enough scientists and the scientific community about the nature of these changes and listened to what was said. Indeed, we have worked very hard at that.

Secondly, the fact is that we do have to make decisions about the design of qualifications, whether it is science or any other subject. That is our role. There is an absolute requirement that, as we look at qualifications, both current and as they change, we do all we can to meet our statutory remit, and our statutory remit requires us to make sure that these qualifications will produce sufficiently valid results. Now, science was not straightforward. It was difficult, because—and many science teachers will recognise this—what is happening at the moment with the current qualifications is that the core, the really exciting bit of science, meaning the practical experience of experimentation, has been moved and shifted over time into a rather stultifying summative assessment at the end. That summative assessment is pretty predictable and it does not produce very different outcomes either. Students and teachers have found that truly exasperating. I have been to any number of schools where I have been in the lab faced with teachers and students on the other side of the bench with their arms folded telling me to do something about it.

There is an absolute recognition that that simply was not working. It was knocking the fun out of science and turning students away from it. Our interest was in getting more practical science over the two years of study so that students really did learn from it. We were not able to square that with an assessment at the end. Not being able to square these things, we have put, if you like, the curriculum intent first. We have said, “If we cannot get to the point where a summative assessment can be done well enough and is manageable enough and produces fair enough results, the answer is to make sure that practical science happens in the classroom. What can we do to make sure that happens over the two years?” It is true that we now have it separately reported, but we have unprecedented checks and balances around that to make sure it happens. Every school is going to be visited over the next two years to make sure it is happening. We are requiring the individual student to keep his own record of the practical science that he does, and we are requiring the head teacher to sign off that those opportunities have been provided for students.

We are not ignoring the assessment at the end, so 15% at least of the marks in the final examination will be allocated to questions that test the student's understanding of experimentation and practical science. If you aspire to good grades, they are not going to be achievable unless schools have given their students sufficient opportunity to really understand practical science. Yes, these things were difficult. I am very thankful of the fact that these changes have been welcomed by a good number of science students and teachers, with science teachers seeing the opportunity to truly integrate practical science into the teaching of science, so keeping that flame of interest burning in young students rather than snuffing it out.

Q26 Kate Hollern: There is obviously strong disagreement, is there not, because evidence to this Committee from the Wellcome Trust stated the pilot you conducted was limited and too short. I do not think you have convinced people there is a robust process in the pilot assessment.

Glenys Stacey: We have, I would say, convinced some people, but not all people. But what we have done is met with the Royal Society, and indeed the Wellcome Trust were represented there, and we have had good discussions with them about both the risks and the benefits of these changes and what controls and measures we can put in to support schools and to make sure, so as far as possible, that these qualifications deliver. Indeed, those discussions were interesting because, yes, they identified changes to the qualification, but also some of the factors affecting whether or not science happens or not in the classroom, which are separate entirely to the qualification. They relate to the resources available within schools and indeed the supply of teachers and so on. The issue is often shoehorned into the qualification, whereas promoting good science and good science teaching is a much wider debate.

Amanda Spielman: Yes, I endorse everything that Glenys has said. Perhaps one further thing: sometimes there is a hope that is often expressed in slightly different ways that the qualifications themselves can somehow guarantee that everything that ought to happen in schools will indeed happen. That is a weight that qualifications on their own cannot bear. It is the responsibility of the education system and everyone in it to make sure that the classroom experience that students get is as it should be. We do sometimes have to point out that perhaps rather more weight is being put on a qualification than it alone can bear.

Chair: Lucy, do you have a quick question on this subject?

Q27 Lucy Allan: Yes, I note what you say. I would simply ask: is it possible to assess practical skills through a written examination?

Glenys Stacey: Yes. It is done internationally in other qualifications and we have seen good examples of that ahead of making these decisions. Indeed, we have seen actual examples, very good examples, in the sample assessment materials that we have accredited at A-level. So the short answer is there is sufficient evidence to suggest that.

Q28 Lucy Allan: You have no regrets about pushing ahead, despite the opposition from many?

Glenys Stacey: As I say, we had to make a very difficult choice here. When we look at other ways of assessing practical science, some of the suggestions that were made to us,

for example, at GCSE, they are simply not manageable. The qualification is a certain size, 120 hours, so if you take half an hour assessing each student and you have 30 students in the class, that is 15 hours out of the available 120 hours, is it not? During those assessments, the other students who are not being assessed are not doing anything. When you think about the practical implications of trying to do something else, given the nature and size of the GCSE cohort for science, it very quickly becomes not feasible. One then has issues also about the credibility of the results, but that is not the primary reason for change.

The reason is to get science in right from day one when you start studying your GCSEs, so you do more experiments, you can happily learn about experimentation, you can do an experiment and make a mistake and learn from that. An experiment can tell you something completely counterintuitive and that is a lightbulb moment. A lot of that, frankly, has been knocked out of science for the last decade and that is simply not good enough for those students. Now, we have looked for the best possible way of squaring these things. It is not comfortable having to make these decisions, but as I say, we believe that we have put the curriculum first.

Chair: Thank you very much. Lucy Frazer, you are going to be talking about comparable outcomes, which is going to be introduced in 2017.

Q29 Lucy Frazer: Ofqual is going to do a comparable approach for the awarding of GCSEs in 2017. If grades are standardised how can the school show progress?

Glenys Stacey: I do not want to give a little tutorial in how comparable outcomes work as I will probably bore you too much, but basically comparable outcomes allows individual differences between the schools. What is set is a national standard, a national expectation, but it does not mean that an individual pupil or indeed an individual school could not improve.

Q30 Lucy Frazer: But if you are only going to get X number of people who get X grade, then if there is a particularly good year nationally then those students will suffer. How will those students get to differentiate themselves?

Glenys Stacey: What happens at the moment is that we apply what I would call comparable outcomes light. What I mean by that is there is an expectation that results will come in broadly in line with predictions based on key stage 2 and other data. When those provisional results come in, if boards have a good case to put to us for the award being higher or lower than is predicted, they put that case to us. This summer, for example, 51 cases were made to us for awarding occasionally below, but mostly above, expectation and we accepted every one of those requests.

Comparable outcomes is not a straitjacket, it is a framework to set expectations and to make sure that results sit sufficiently well with the statistical evidence that has always driven results over many decades. What we are saying for 2017 is that we will apply comparable outcomes strictly, so that we will expect results to match expectations based on key stage 2 and other statistical evaluation. We are doing that because we know from history that when new qualifications come in, for that first year we can expect two things. One is that teachers are still getting used to the new qualification and the content, we can expect them not to do as well as they would do in subsequent years in teaching it. If we left it as a freefall, results

would fall, and those first students in that first year would suffer unfairly because of that. There is an ethical imperative to apply a strict comparable outcomes approach in that first year to protect those students. That is why we do it.

Q31 Lucy Frazer: You are putting a huge amount of emphasis then on achievement at key stage 2 because that will, on what you say, massively influence the grade that the schools are getting later on.

Glenys Stacey: I wonder, Chairman, whether I can suggest we do a note to the Committee on how comparable outcomes work, because I think with a good number of new members that might be helpful. But in summary, all of the research shows that key stage 2 is a good national predictor of achievement at GCSE, not an individual predictor. It will not tell me how well or badly you would do or I would do, but as a national predictor for the cohort it is shown, through very good research, to be reliable. It has been relied on as one of the key statistical datasets for awarding for many years in exam boards. It does not mean that we cannot get better at awarding and we are always looking for ways to improve awarding. Indeed, you will know that we are looking at the benefits and drawbacks of a national reference test, for example, to help us as well. But at the moment it is a very well-established and respected aspect of awarding.

Q32 Lucy Frazer: Will you be doing comparable outcomes post-2017?

Glenys Stacey: We will be doing comparable outcomes post-2017. There is a debate to be had and consideration to be had of how purist an approach we adopt to that: will there be tolerances; if so, how much? That will depend on judgments we make later on particularly around the results in 2017. How stable does the picture look? Is it then the right time to introduce some tolerances and allowances or should it still be strict outcomes? We will be making those judgments openly and transparently with the interests of standards and students in mind.

If I might say one other thing about 2017, 2018 and 2019, and AS of 2016, we know that, when changes of this nature are brought into qualifications, results are more volatile school by school. We can hold the national standard, and we will, but we can expect, quite properly, that some schools will be better than others at understanding subject by subject the detailed differences, for example, in the curriculum. We know—we have seen it every time—that there will be greater volatility, greater variance among schools, and that will happen for a period of time as we go through these changes. What we would suggest is that that needs to be understood. That does not suggest that the qualifications have failed or that schools or teachers have failed. It requires intelligent interpretation to make sure that these changes are understood and indeed that they are expected. This is how implementation plays out in practice. It is helpful to be aware that does not mean that there is something wrong.

Amanda Spielman: We have published and will continue to publish information on that.

Q33 Lucy Frazer: Can I ask you a question about the 1 to 9, A to C? Obviously if you go from A to C to 9 to 5, you are increasing the number of levels and therefore making distinguishing between students easier. Is there a case for not having grades at all, but just

having numbers? Then it is easy to see and you are not stuck on, “Did you get a 6 or a 7” but it would be 86% or 90%?

Glenys Stacey: There is an argument to be put for moving away from grades altogether and just going to raw marks or some sort of standardised mark. You are quite right. That has the benefit of not drawing what one might say to be artificial lines to differentiate. There are pros and cons to that. For example, it places a relative certainty around the mark, whereas there is always an element of discretion in marking. Grading is designed to try to manage that, to try to get some certainty. Indeed, I think research has shown that, if you have about eight or nine grades in a qualification of this sort of size, your prospect of being wrongly graded is fairly small and your prospect of being wrongly graded by more than one grade is very, very small indeed. There is a balance to be struck, as ever. There are always trade-offs in the design of qualifications and so there is a case to be made for grading in terms of you can be pretty sure that the grade is right.

But there are benefits to not grading and having a mark, particularly if you can then show how certain you can be about that mark and if you can put some tolerance bars around it. Now, that is a very different approach. It is not something we have done in this country. It is done in some other jurisdictions and the research I have seen on how it is regarded there is a bit mixed as well. Maybe that is a debate to be had next time that we do a round of reform.

Amanda Spielman: If I can add, I think it was one of the things that we consulted on. We were open to considering this.

On the point that was made earlier about our responsiveness to reactions, there was a very clear reaction from schools and colleges, who have a strong preference for a graded system. That is my recollection. There was a reluctance to embrace change in that direction as well as in the content of qualifications.

Q34 Caroline Nokes: I wanted to ask a bit about the national reference test coming in in 2017 and specifically how are the schools that are going to take part being selected? Drilling down into that, how are the pupils within those schools—the lucky or unlucky 60, depending your perspective—going to be selected, and will any school or any pupil or parent have the ability to opt out?

Glenys Stacey: A few things to cover there. First of all, we have a contract with NFER, a well-respected organisation, very good connections with schools, and it is their responsibility to select. We have done a fair amount of work with them specifying how that is to be done and I am happy to send you a note about that, Chair, so you can see the detail. But in short, we are looking to secure a truly representative sample of schools and of students. We have done the first round of trialling a couple of weeks ago and did see a wide range of schools approached in that.

You have asked about, I think, protecting students. Again, we have, in discussions with ASCL and then work with NFER, made sure that we are able to make decisions as to whether there is a particular case to exclude any particular students or pupils from having to take the test. There is some discretion there that we would encourage the use of to make sure that, as much as anything, the test produces good results—fair, reliable results—but also so that students are not unduly pressured. If they are thought to be unduly affected by the prospect of a test, for example, we would wish to have that properly considered.

Have I answered all of your questions?

Q35 Caroline Nokes: Would a parent have the ability to refuse to allow their child to take part?

Glenys Stacey: Not as far as I know. It looks like it is head teacher's discretion. Yes, head teachers have the option to exclude pupils from the testing and no doubt parental pressure would be something that the head teacher would consider.

Q36 Caroline Nokes: Do you have any concerns, given that year 11 is already a pretty pressured time for teenagers, that this is loading yet more pressure on some of them?

Glenys Stacey: This is not the vast majority of students that are taking it; it is just a small sample. Secondly, they are not being asked to take the entire range of papers. They are taking selected papers, so the actual demand is for a short length of time. It is not a full assessment and it is not a full GCSE assessment. It will be one of six papers, for example, so it is relatively short. Of course I am aware that this is a stressful time. Again, this is a balance that we are trying to strike between the imposition that we place on schools and students on the one hand and the potential value of a reference test for education as a whole and the confidence it could build in results on the other.

We think that is a balance we have to keep looking at and keep reviewing, so we are not sitting here with our colours nailed to the mast of the national reference test come what may. We will continue to consider the benefits of the test as against the imposition and indeed the cost on the system that is there. At the moment we think this is the right thing to do and worth taking through the pilot, which we are doing now, but we will always keep it under active review. I can assure the Chairman that my board has a very keen interest in that balance, so do not assume that this is going to happen come what may. It is about evaluating all the time the benefits as opposed to the potential drawbacks.

Q37 Caroline Nokes: Statements made by Ofqual suggested that the tests are going to be broadly similar from year to year and that there will not be much variation. The teenager is an enterprising beast. Are you remotely alarmed at the prospect of the papers ending up on the internet for next year's cohort to be able to download the answers and just whizz it as you might a spelling test?

Glenys Stacey: Just to be clear, we know from how this style of test has been developed and used in the few other jurisdictions that have done it—Hong Kong, for example—that it is important to keep the stability of the test. If you change the test items substantially, you lose the ability to compare it with sufficient granularity to make the test meaningful. It is a fairly fundamental thing.

Secondly, NFER have very clear commitments to keep control of the test. It is not that the student will be allowed to see it, take it, copy it or whatever. But, yes, we are aware, of course, how enterprising students are and we do think there is a risk that there will be some familiarity with the test. Our students will not know until a very short time before that they have been selected to take it. There are some reasonable checks and balances that we have put in there to try and manage that risk, but we will have to keep a sharp eye on it, it is true.

Q38 Caroline Nokes: Can I just take you back there? Teenagers will find out a short time before they sit the test. How short is short?

Glenys Stacey: Do we know? Two months.

Q39 Caroline Nokes: That is not short at all, is it?

Glenys Stacey: Do not forget that we do keep an eye ourselves on social media. We are not unsighted in relation to what is exchanged on social media. We will be tracking that and so will NFER. But the point I am making is the student is not sitting there a year ahead looking for the paper to prepare for it.

Q40 Caroline Nokes: There are two months. I would have said that 24 hours would give an enterprising kid enough time to go and have a quick Google and see if they can find it.

Glenys Stacey: You will appreciate that 24 hours will not allow the head teacher to consider—

Q41 Caroline Nokes: No. That was my point, that 24 hours also would not give a parent time, who did not want their child to sit their mock GCSEs, sit a national reference test and then go on to sit their GCSEs in the space of four months.

Glenys Stacey: That is right. We do have to give reasonable notice. There are the normal procedures to go through, and indeed for schools to be ready, and two months seems about right. That is the judgment that has been made.

But yes, I would like to make it plain that we know there are risks in the national reference test as well as benefits. We are carefully and actively evaluating those risks and indeed the costs of the tests, and if we think that the balance changes, then we will think more fundamentally about whether that should proceed. But for the moment there is a strong desire with those in schools particularly that there is some objective measure other than awarding as to whether or not teaching and indeed learning is improving over time in schools. There is no other objective measure, so it has a strong potential to play its part, but it is not without problems, I know.

Amanda Spielman: There is no advantage to the individual member in looking it up, because they do not get a result for this test and the school gets no result for this test. The point is to make sure that people do not have incentives to try to prepare for it. It is just something that takes an hour of their time. It should be a sort of passing distraction in the wind almost for the teenager. This is not—

Q42 Caroline Nokes: I think it is a very difficult time for a teenager to have a passing distraction when there is already so much pressure on them in the build-up to their GCSEs. Do you think it is fair—I have a rather cynical view about this—that, if you had confidence in GCSEs, you would not need this?

Glenys Stacey: I think that is a little bit harsh. My own view is that the way awarding is approached in this country is impressive. I do not know of another jurisdiction that does it better. There is a lot of earnest work done in exam boards to make sure that their

prediction matrices are as complete as they can be and as good as they can be, that examiners do a good job, that they keep that precious kernel of awarding safe from any undue influence and that they are as confident as possible of those awards. It is also true that, as a regulator, Ofqual oversees that awarding very carefully and sets expectations. Indeed, we have set them already for next year.

We meet in formal meetings with all the exam boards and the other regulators across the UK to keep track of that awarding, to review provisional results, to hear of any concerns and so on. This is a good established process of awarding, but schools want to know whether there is any other objective measure. That is what the national reference test is designed to produce but, as I say, it is not without difficulties and it certainly does not suggest, and should not be read to suggest, that we have a lack of confidence in that awarding process in this country—not at all.

Q43 Kate Osamor: This is to Amanda. There was a recent conference where you spoke about a study around re-marking and it revealed that markers were more generous basically straight after summer. We just wanted to know if this significant. Is it epidemic and is there an issue that we need to know about?

Glenys Stacey: I should answer that because I am more up-to-date with the information. What the study has shown is that, if you take markers and ask them to review marking three months after rather than immediately after, which generally happens, it shows that markers are much more inclined to change the mark and they are more inclined to lower the mark than to increase it. What that seems to tell us is that, when you are more immediately reviewing marking, the human factor is more prevalent so you are more likely to be aware, for example, of how close to the boundary the mark is and what is riding on it for the student. That is because markers are examiners. They are human beings.

Amanda Spielman: They are teachers.

Glenys Stacey: And they are teachers, so it does not surprise us to find that if a length of time goes by and that immediacy is gone that the judgments are slightly different. It is helpful to know that.

Q44 Kate Osamor: But there seems to be such a wild variation from the initial mark to the mark afterwards. With that, it is causing lots of concern within schools and within the teachers. How are you able to ensure that there is consistency, because if the variation is so different, it almost puts the first mark into disrepute?

Glenys Stacey: Yes. Just a few facts to get across then. Last year there were 400,000 requests for review marking. Four out of five of those requests resulted in no change to the mark, then across the spectrum of results as a whole, less than 1% of grades are changed. Our research has shown that most appeals are happening between just one, two or three marks of the grade boundary, the upper grade boundary. It is not an even spread across the marks. It is when it is very near the grade boundary and you can understand why an appeal would be particularly appealing, and then secondly that the average change is very small indeed. In fact, I recollect the average change is about one mark—one mark can take you

over the grade boundary. This notion that changes are routinely of a more significant size is plainly wrong.

But what we do know is that, last year, for example, of the many millions of scripts that were marked, 627 requests resulted in a change of more than two grades, so what is going on there? We have looked at that and what we know is going on is that, contrary to popular belief, it is not necessarily rogue markers doing something completely unconscionable. It is generally systems issues, so perhaps the script has been disaggregated so that individual questions are marked and then it is not put back together fully, or put in simple terms, the adding up of the marks from each of the questions is not right or is not complete. The issues for exam boards are often about systems and improving their systems so that they are as foolproof and robust as possible.

It is also true to say that, with 50,000-odd markers each year, there will be occasions where markers are not working well enough. Exam boards are getting increasingly good at developing and applying those systems to identify where marking is not good enough in the live environment. Now that much more marking is done electronically, exam boards can supervise that live. They can do what we call probity checks, so if you are marking and when “Coronation Street” is on, your marking is “three not, three not, three not, three not” an exam board can identify that and wonder whether you are marking with quite the same attention as you would when “Coronation Street” is not on.

Kate Osamor: I do not know if we can blame “Coronation Street”!

Chair: It says quite a lot about “Coronation Street”, does it not?

Kate Osamor: I know you are using it as an example, but what is important is that the grades are correctly marked and whoever is doing the marking understands that this is crucial to the student and crucial to the school as well, because it has a major impact on the school. If a school has small resources or limited resources, how are they able to appeal effectively, because what will happen the next year? It will deter certain parents sending their child to that school.

Glenys Stacey: Absolutely, yes.

Kate Osamor: So it is very important for yourselves as the regulator to ensure that you have the best markers that are not watching “Coronation Street”. But on a very serious level, that is your job.

Glenys Stacey: Absolutely. Just to be clear about our job, we regulate exam boards and we require exam boards to assure us of the quality of their marking. We do not recruit the marker workforce or indeed evaluate it directly, so it is important that our respective responsibilities are understood. But yes, we want assurance and we also want to know that marking is of good enough quality. We did a seminal piece of work and published a full report in February 2014, a report welcomed by this Committee, which has research showing—counterintuitively for most—that our examiners and our markers are generally very experienced in their subject, often not only graduates but having post-graduate qualifications, for example, and many of them are heads of subject and have been marking for many years.

But nevertheless there are always improvements that can be made to the quality of the marking and exam boards will admit nowadays, thankfully, of particular problems in certain subjects, for example. We know that for each of those exam boards, we have required them to agree with us an action plan as to how they improve their marking. Those action plans are tailored to the exam board and the subjects on the particular issues that they have. We are monitoring those action plans to make sure that they are being delivered.

Coming on to your second point about the fairness of the appeal system, absolutely so. We want a system that is swift, effective, fair and has public confidence. We recognise issues with the current system, so when you speak of whether one candidate or another can afford an appeal, what we do know, for example, is that some schools uniformly appeal within one or two marks of the grade boundary, come what may. I would call that strategic appealing. That does not seem to me particularly fair to any individual student either in that school or in another school, so there is a fair amount to consider when it comes to fairness of access to these systems. But certainly the cost of it is one consideration. We are very keen on making sure that the cost is contained and justified as well.

But we are about to make some propositions—I know that you are aware of this because we met the other day—to, we think, significantly improve the system of appeals and I think we will get those propositions out towards the end of the month or turn of the month. They are based on a good amount of solid research we have done as to how it can be improved. I hope that you will welcome that new research.

Q45 Ian Mearns: I am just wondering, we have had evidence from SCHOOLS NorthEast, which is a representative body that covers about 1,200 schools in the north-east of England. The evidence that they have submitted has gained resonance and has been featured in the regional press in the north-east of England. Do you think it is worth your while having a conversation with SCHOOLS NorthEast about their evidence.

Glenys Stacey: I think it is. In fact, I was going to suggest that we would very much welcome any communication from that body. We have not heard from them as yet, but of course we would welcome that.

Q46 Ian Mearns: Because their evidence is quite punchy in many respects and they have described the volatility in performance of exam boards as alarming. They say it has greatly undermined confidence in the system of assessment. Of course, they would argue, I think probably quite rightly, that that has a profound impact at the human level when you look at individual students. When you talk about the appeals process being swift and effective, part of their submission suggested that one school were told that their grades would not be altered within two hours. That is swift, but is it effective?

Glenys Stacey: You have raised a number of things there. First of all, in terms of volatility among schools, we published on the day of results for A-level, and indeed for GCSE, some volatility information—they are able to show it nowadays because of some statistical analysis in the day or two beforehand—the extent of volatility and how it compares with volatility within subjects this year and previous years, so I will be able to refer SCHOOLS NorthEast to that and we can see how schools in that area sit.

Secondly, I think it is right to just confirm that, as we have moved to linear qualifications, awarding is done wholesale at the end of two years with a good amount of data and information about how students have performed across the whole of the qualification and the assessment. I think there is significantly more confidence in those awards. This year, for example, when I have met with exam boards alongside other regulators to review provisional results, there have been no concerns expressed by the regulators or by the exam boards as to any individual award. There is a growing confidence that has been induced by the move to linearisation that we can have in the awards that we now see.

Of course, you will have seen the national picture. I would hope that, in recent years, given the work done at Ofqual and in exam boards to truly be confident about awarding, the public can gradually have increased confidence in those awards as well. But of course individual schools are affected year by year. Many things affect schools: changes in teachers, changes in choice of one exam board or another, and indeed changes in the nature of the cohort at any one year at school level as well. These things are often hard to evaluate when results do not come out quite as the school might expect or hope, but we are always happy to talk to schools about their results.

Q47 Ian Mearns: I am sure that SCHOOLS NorthEast would welcome the opportunity to have a conversation with you. I suppose when they are using terms like “exam board roulette” in terms of grading and assessment, they have a story to tell, I would suggest. I will be in touch with Mike Parker at SCHOOLS NorthEast and ask him to give you a call.

Glenys Stacey: Yes. The term “exam board roulette” I have not heard before and it is certainly one I would like to debate.

Chair: I think Mike Parker might be watching this. I happen to know he is interested.

Glenys Stacey: Good. I look forward to meeting him.

Q48 Michelle Donelan: I just wanted to touch on the crisis that happened last year that led to OCR failing to release GCSE and A-levels on time. Well, almost failing.

Glenys Stacey: Succeeding, I think.

Michelle Donelan: Do you think that that is symptomatic of the exam system, that it symbolises that it is under far too much pressure, and that there are not sufficient resources?

Glenys Stacey: The position we found ourselves in with OCR last year was that, ahead of results, we became aware that they were struggling to keep up, if you like, with their marking. They were not making sufficient progress on a daily basis, so we increased our interest and oversight of their marking and were having daily meetings with them and daily exchanges of information, subject by subject, plotting progress and making sure that they were applying sufficient management oversight to make sure that results came in on time. The results did come in on time and we subsequently produced a full investigation report, published earlier this year in the summer, detailing the exact circumstances that arose there and what went right and what went wrong. In summary, this was a matter of noted management failings within one exam board, so it is important to recognise that this was not about marker supply or marker competence. It was about the management of the marking process by one particular exam board.

Q49 Michelle Donelan: But then at the same time, Cambridge Assessment claims that Ofqual is making exam boards into political scapegoats and using them as an excuse. Do you not think that it is more endemic—that the actual system that is failing rather than particular exam boards?

Glenys Stacey: As I have said, in 2014 OCR—and OCR do recognise this—did not manage their planning or delivery of marking well enough. Other exam boards did in 2013, 2014 and 2015, and indeed OCR managed their marking sufficiently well this year as well, having learned from the mistakes they made in 2014. I stand by our investigation report.

Q50 Michelle Donelan: What steps are you taking to alleviate the shortage in well-qualified examiners? We touched on that a little bit before.

Glenys Stacey: I do not know what the evidence is of a shortage in well-qualified examiners. As I say, all marking was delivered in good time this year and the marker/examiner workforce, as far as I am aware, is sustained. Of course we have moved to linear qualifications. In the past, the system was supporting marking of many units and so the demand has changed as well, but I am not aware that there was a marker shortage this year. I think we just need to be careful about us assuming that, when the evidence is that marking was done.

That is not to say that exam boards should not be always thinking ahead and always preparing for the particular challenges. Yes, qualifications are changing, mark schemes are changing and the balance of types of question are changing within assessments. Exam boards are looking at that, looking forward to that and always looking to recruit and keep refreshing their marker pool. They are doing what seems to be some good work with HMC and ASCL to encourage good teachers to get involved in marking. We would always encourage that. It seems to us, put simply, 52,000 people predominantly in schools at the moment—good teachers—play their part in marking. The more schools and teachers can respect and play their full part in that important part of awarding the better. It is in our collective interest that people do that and schools respect that and find a place to applaud that as well. The more that can be understood the better.

Q51 Michelle Donelan: You mentioned the change in assessment from module. Will that not potentially cause a problem because you will need more examiners at one time?

Glenys Stacey: Yes. It changes when the demands occur.

Michelle Donelan: Yes, because this is all at once.

Glenys Stacey: That is right. Exam boards are acutely aware of that and of course they have managed that this year. We have had linear qualifications and it has been managed. What I am curious about is changes not just in the style of assessment—one can often over-exaggerate the impact that that will have on markers—but in changes in the take-up of qualifications. As we see the move to “best eight”, for example, what we might predict is that, on average, students might take a smaller number of qualifications than they have done, so perhaps the days when students routinely take 12 or 14 GCSEs may be on the wane. We may see increases in take-up in some of the core subjects. Exam boards are

trying, through a glass darkly, to predict these changes and to make sure that they have markers in the right subjects and the right numbers of markers. That is all a little bit difficult when, for example, take-up, yes, is pretty uncertain.

Q52 Chair: Do you have any data on the turnover of markers?

Glenys Stacey: We certainly have information from the survey of markers to support our research in February 2014. I think 32,000 markers—as I understood it, I will check—responded to that survey. We produced information that shows that the majority of them have been marking for many years. They are not wet behind the ears. A big majority have been doing it for more than 10 years and many were heads of subjects, so it is a more stable workforce than one might think. I have said this before and say it again: markers do not generally do it for the money—not surprising, given the money—but they do it and they tell us, and told us in the survey, that they do it predominantly to keep close to the specifications in the marks, to keep right on top of it, to truly understand those qualifications so that they can teach them well and that their students will do well enough as they go through the assessment. There is a sort of win-win in a sense. They get something out of it, and particularly so when a new qualification comes along. There is a great incentive to get under the skin of it and understand it and one way of doing that is by becoming a marker or examiner in it.

Q53 Chair: A good bit of professional development, in fact?

Glenys Stacey: Oh, absolutely. From what I can see, it is much more embedded as a fundamental part of development in some other countries. We seem somehow to have come to a place where it is almost an adjunct. We would like to see it more central and more well-regarded for the contribution it makes to the integrity of awarding. Marking is just as important as the actual awarding and the grade boundary setting. It has to rely on good markers to get to that point.

Q54 Lucy Allan: I wanted to ask about the new regulated qualifications framework, which has broadly been very well-received because of increased flexibility and greater focus on vocational qualifications. Following up from that, I just wondered if you could comment a little bit on comments from the Association of Colleges, which has suggested that, with 60% of 16 to 18 year-olds following a vocational programme, Ofqual should focus more heavily on vocational qualifications.

Glenys Stacey: I am so glad you have asked me that. There is so much attention on general qualifications and I absolutely understand why, but most of our awarding bodies are in a different market. What we have done, about 18 months ago we recognised that we could raise our game on vocational qualifications, so although it is rather tedious to those outside Ofqual, we have restructured the whole organisation, created a vocational qualifications directorate. We have an experienced executive director for the vocational qualifications, Jeremy Benson, and we have, on the last count, about 50 staff now dedicated in that area.

Amanda Spielman: I can add we have also established a vocational advisory group to the board, which I chair, which provides invaluable input to making sure that we think about this.

Glenys Stacey: What have we achieved so far with that? First of all, we have developed a very good understanding of the diverse nature of so-called vocational qualifications.

Secondly, we have taken steps to review functional skills qualifications, where there are over 1 million and rising candidature. They are significant qualifications within apprenticeships and elsewhere, and particularly important in colleges, so revising and improving those qualifications. We have done comprehensive audits of a good group of vocational qualifications providers, focusing on particular qualifications that are thought to equate to others but perhaps do not, so making sure that we get to that.

We have unpicked the QCF. It is quite interesting, people generally deride the QCF, and I understand why and would probably join them in that, but when you create something like QCF, people do not think about, “Could we readily unpick it?” and unpicking it has been extremely technical. We have done it, we have achieved it, and it is a small but not insignificant step for mankind in getting to the place where we do not have the QCF. We are making quiet inroads into an area of work that is not often focused on, but we think is significant to such an extent that despite the pressures of general qualifications reform we dedicate a third of our resource to it.

Chair: Ian, you are next, on the role of Ofqual regs.

Q55 Ian Austin: Over the past five years, Ofqual has had to work closely with the Government to reform the qualification system. Do you think Ofqual and Government might have become too close?

Amanda Spielman: The short answer is I do not think they have. It is not an easy balance to strike, being an independent regulator, making sure that you work constructively in the areas you have to work constructively, but at the same time you remain able to take independent decisions and given independent advice. There are plenty of examples we can point to where we have made difficult decisions where there are competing interests in different directions and where we have given good advice to Ministers. Perhaps delaying the GCSEs other than English and maths, or rather, advising Ministers to delay other GCSEs was one example of that. Having said that, there are others to point to.

Glenys Stacey: There are other examples, for example, advising Government to postpone new reform of A-level maths for a year so that we have consistency between GCSE maths and A-level maths. Given the nature of the changes in maths, that seemed entirely right. Taking the view that we have taken in the sciences would be another example. There are difficult decisions we have had to make, but we are very clear about our role and the statutory objectives that we need to fulfil.

Q56 Ian Austin: Cambridge Assessment described your relationship with exam boards as, “legalistic and prescriptive”. They claim that this stifles innovation and has led to a regulatory approach that is based on command and control. Why is Ofqual’s relationship with exam boards become so strained from that perspective?

Glenys Stacey: I do not know that it is as strained as it is perhaps portrayed there, but if I explain that we regulate the providers of GCSEs, AS and A levels differently to other providers in some respects, so it is necessary, in fact, to be prescriptive about some of the

requirements for GCSEs. If we were not, you would be hauling me over the coals for allowing standards to race to the bottom. It seems to us that if we are to have a model with four exam boards or more and we have prescriptive content from Government for these qualifications and an expectation that within each subject these specifications will be truly comparable, then we need to be pretty prescriptive ourselves about how that is to be achieved, but it is not quite command and control.

The way that these requirements are formulated is that we meet with exam boards on a regular basis, monthly, generally whole-day meetings where the chief executives of those boards and their seconds in command come along. We have an agreed agenda. What we are doing there is talking through detailed design issues for these qualifications and particularly how things would affect the manageability and the impact in schools, what is the best way, for example, to make sure that any particular non-exam assessment is well enough moderated would be a pretty good example. We discuss these things together. We do try to agree an agreed approach to the best thing to do in these issues, which is then about embodying, if you like, an agreed notion of quality. Yes, sometimes this is about swapping the adverse impacts of competition while promoting the positive things, and those agreed requirements are then embodied in regulations that we then require exam boards to meet.

Often as not these are agreed between us; not always. Sometimes, as you would expect, the regulator will take a different view. Often as not exam boards do not agree among themselves as to what the right approach is. Why should they? If we take as one example should there or should there not be a requirement for a minimum assessment time in maths GCSE, there will be different views about that. We regulate it for a minimum time because otherwise there might be a race to minimum time.

It is fair to say that there is an unprecedented amount of regulation on these qualifications. That is no bad thing if what you expect from it is good qualifications that deliver the best educational outcomes that you can have confidence in and they are truly comparable. Yes, regulation is frustratingly legalistic. It is, because it has to be. A good regulator needs to be able to use the whole gamut of their powers, including the powers of enforcement and sanction, and to do that with ease and authority. You only are able to do that if you have started right at the beginning on the legal basis upon which you are doing anything. If you can do all of that with ease, then you very rarely have to use those powers, just a conversation will achieve what is to be achieved. That is the way regulation is.

Q57 Ian Austin: From your perspective, this strident view of Cambridge Assessment is not going to lead to there just being three exam boards?

Glenys Stacey: I think it is for exam boards to decide whether or not they stay in the market and I hope that they all do.

Q58 Ian Mearns: I think this was also reflected in the evidence given by SCHOOLS NorthEast, that the perceptions survey has highlighted a deterioration in trust in Ofqual from teachers and head teachers. Is there anything that you want to do about that?

Glenys Stacey: It is important to understand why that might be. It is difficult to tell from a perceptions survey, but from conversations with those in schools we get some sense of it. Certainly trust took a big dent in 2012 because of GCS English, a qualification itself that

sort of over, bent under pressure. Secondly, I think there is a confusion of roles, what does Ofqual do and not do, which perhaps does not help, and then thirdly, as the accountability pressures bear evermore on schools, people are understandably evermore anxious about their results. There is no doubt that as these reforms come to fruition, some of the weaknesses in the old qualifications are being removed and that does make it tough.

Q59 Ian Mearns: Do you think, given the result of the perception survey, that there is a job therefore for you to do some follow-up work on that? The Secretary of State told us a couple of weeks ago that she has spoken to 1,000 teachers since she came into post. Do you think there is an opportunity for Ofqual therefore to be doing some more rigorous survey work directly to teachers and head teachers to get to the bottom of this?

Glenys Stacey: Yes. There are a number of things that we have done. For example, we ran an ethical survey for teachers and they were able to express many different views as to what they thought was right and wrong, the way qualifications play out in schools. We have had a day symposium with teachers and their representatives at Oxford University, with Oxford University, talking those things through in a non-threatening way, understanding absolutely the pressures on schools.

Thirdly, we get out to a good number of schools. There is no doubt at all we are always trying to get out to schools and we always bring things back from that. I think it is one of the successes I would claim for the organisation, that we do not pretend the world is different to what it is, that we always encourage absolute candour in a non-judgmental way, being absolutely sure we understand the real world in schools in every sense. Just to finish, without that understanding, we would not be able to design qualifications well enough.

Amanda Spielman: Following through on that, I think we are interviewing today for a teacher fellow to come and work at Ofqual for a year and help make sure that we are addressing the right questions and thinking about the right things.

Q60 Ian Mearns: I am not suggesting that there is some sort of cosmetic exercise that needs to be done here, but I think it may be useful to the profession out there if there is clear sign-off, a tangible sign of engagement of that nature and which everybody can see in an open and transparent way.

Glenys Stacey: Thank you. As I say, we do do a good amount of engagement, but there is always more to be done. What strikes me most in speaking with schools about this area is just an understandable belief that we are responsible for everything. It is not quite like that.

Amanda Spielman: I very recently visited the school where Suella is chair of governance, for example.

Chair: Lucy, would you like to talk about national assessment levels?

Q61 Lucy Frazer: Yes. The Government recently removed the curriculum levels, allowing schools to make their own assessments, and Ofqual took part in the commission on assessment about levels. Do you think that the Commission's report adequately supports schools in their assessments?

Glenys Stacey: I will just make one point, and that is we were there to provide technical advice to the Commission and we did that. But would you like to deal with it?

Amanda Spielman: The Commission report very significantly stopped short of providing specific advice on how schools should do it. It gave them a number of principles and steers. I think that there is still a considerable amount of discussion about the next steps for translating that into specific practicalities for different types of school. That is not something in which we have a direct role, but again, we stand ready to provide advice as and where it is necessary and helpful.

On national assessment more generally, our responsibility, as in other parts of the system, is to make sure that there are good yardsticks that do not stretch and bend. We are very much wanting to make sure that what is used by Government to measure through the various kinds of testing check that it operates work for the purposes that they are intended.

Q62 Lucy Frazer: Although it is not your primary role, will you take part in monitoring the reliability of the tests when they have been worked through?

Glenys Stacey: Absolutely so. That is our role, making sure that those national assessments are sufficiently valid, that the results can be relied on. Indeed, we wrote to the Select Committee a few months ago now to explain that we were moving our focus from procedural risks, delivery risks, now that we have an executive agency to the STA response for these tests to focus more on how valid are these tests.

Chair: Thank you. Suella, would you like to say a few words about overseas qualifications?

Q63 Suella Fernandes: Yes. There was written evidence submitted by Cambridge International Examinations and International Baccalaureate Organisation criticising the extent to which you are regulating overseas qualifications and asking for there to be more clarity within your remit. What is your comment on that?

Glenys Stacey: We do what we are required to do by statute and the statutory provisions require us to regulate those qualifications on offer in England, so if the qualification is also on offer in another country that does not absolve us of our responsibilities.

Q64 Suella Fernandes: Yes, that is where the grey area is, because although the 2009 Act does say that you are to regulate qualifications in England, there is dispute as to whether that should cover regulated business overseas and that is where there is challenge.

Glenys Stacey: We obviously have taken due legal advice as to our responsibilities and believe we are absolutely clear and sure about them.

Q65 Suella Fernandes: What is your view of the EBac and the impact that that will have on standards and your regulatory ideals?

Amanda Spielman: Lucy mentioned a good pass earlier. Like that, the EBac is very much a matter for Government. It is a performance measure; it is a conversation between the DfE and schools. It is like any performance measure, any lever that is applied in school accountability will have some knock-on effects in the exam system, so we have to be

aware of the places where strain is possible and to make sure that our regulation monitors and addresses that.

Q66 Chair: We have had some written evidence that we will be publishing, obviously, that suggests that there is concern that the changes to GCSE and A levels is going to be difficult to complete. Do you think that is a legitimate concern?

Glenys Stacey: Difficult to complete?

Chair: Yes, so a lack of confidence, basically, in Ofqual's ability to complete the reforms is essentially the thrust of some of the written evidence. Are you confident?

Glenys Stacey: Yes. If you remember, Chairman, a couple of years ago your Committee recommended that we increase our capacity and indeed competence and we have been working assiduously to do that. Our numbers have increased. I am slightly embarrassed to say that, given the pressures elsewhere in the public sector, but we have put a justifiable case for that and we have done that. Not only have our numbers increased, but the competence of our staff. We have been training our staff, we have been selecting good people to come and join us as well, including particularly looking for expertise in assessment and in programme management. That is not to say that reform at the pace we are going is anything other than eye-watering. It is an eye-watering pace, but we are keeping up with it. If we thought there was any risk to the timetable, we would simply divert more resources to it.

Amanda Spielman: The board does not have concerns about delivery of the programme on time.

Q67 Chair: Good, good. I just had to ask that question, because we have had comments about it and I wanted to give you an opportunity to answer it.

Given there is talk about a single examination board, have either of you got comments about that kind of discussion and debate?

Amanda Spielman: I think it is very important to be very clear that this is a matter for Government policy, not for us to have opinions. Our job is to work with whatever system is set for national examinations and assessments, that we will do what is necessary to make any system work. When we wrote to Government nearly three years ago now in the context of the proposals that were being considered at the time, it was to express concerns about the risk of exam board reform and full-on qualification reform simultaneously. As a Government looking forward now to the future, that is very much a matter for Government. We provide information when we are asked. At some point, from the discussions we have had, we expect to be asked to advise, but we have not done that yet.

Glenys Stacey: If I may say, the Select Committee itself has done good work evaluating models and produced a full report just a couple of years ago.

Chair: Yes, we did.

Ian Mearns: That is the right answer.

Amanda Spielman: I have pointed a number of people to that when they ask for analysis of the subject.

Q68 Chair: Yes, because we will be thinking about this issue obviously in due course as the debate rumbles on, but we have already put down a marker, as you have just said.

Glenys, I want to thank you very much for all that you have done, because it may be that you do not appear before this Committee again. You have always come with remarkably good, clear answers and I want to express our appreciation for that.

Glenys Stacey: Thank you.

Q69 Chair: I was just wondering what you thought your legacy might be once you leave the position you are currently occupying.

Glenys Stacey: Thank you. I achieve precious little on my own. I have an excellent team of people in Coventry, but I do not do the real work, and on their behalf I would certainly be proud of the work we have done to oversee awarding and to make sure that we can have confidence in those awards. We have seen that reflected in a steadying-off of grade inflation, and I hope over time—and these things do take time—that public confidence will increase because of that. I also think that together we have established Ofqual, which was a fledgling when I started, and it is now seen as part and parcel of the fabric of education and is seen as an independent and mildly competent regulator.

Then thirdly—I hope this continues, and I am sure it will under Amanda’s stewardship—that the efforts we have put in to having honest conversations, well-informed, well-evidenced, but honest real-world conversations about how things are, I hope that that approach continues. I do believe that regulators, to succeed long-term, always need to be authentic and that is what we have hopefully achieved so far and it should continue.

Amanda Spielman: I would endorse all of those and I am very glad to hear you say what you have said, Neil, because I think Glenys has fully earned it.

Chair: Thank you both very much indeed for coming along today. Thank you.