

Cambridge Assessment Conference 2012

Glenys Stacey



Good morning.

When Paul Newton asked me sometime ago if I would take part in today's conference, I was pleased to say yes. That was partly because it was Paul Newton (who would say no to someone so eminent in the field of assessment) and partly because of the subject, risk.

When regulators get together, that's pretty much all we talk about – risk, and the risky things we regulate. It is always a hot topic for us.

Since Paul asked, we've had proposals on A level reform, and on GCSE reform or replacement, and the summer exam series. They are all relevant to risk, but the views I express today are not shaped by those things, but simply brought into sharp relief. The concept of public confidence, and its relationship with risk, is longstanding.

I know that many of you here work in qualifications and you've been doing so for many years. I hesitate to talk to you about the challenges to public confidence in qualifications, given the depth of experience here, but I shall do my best.

I'll start by discussing public attitudes to risk, and public attitudes to qualifications. I'll then look at how things stand – what information we have on public confidence in qualifications and then go on to look at risk in three areas: reform, delivery and standards.

Public attitudes to risk first of all.

We know from research that people generally accept that in a large scale, complex system, like the qualifications system, not everything works perfectly all the time. Sometimes, things don't work as well as they should. People understand that. But it may be much harder to think of it in risk terms – to accept the risk that things happen. And most people find it a challenge to make judgements about risk – to decide, for example, what is an acceptable risk. Risk is alright in theory, but in practice, individual risks are extremely uncomfortable for people.

And we have a rather disjointed view, as people, about how risks should be managed. On the one hand, we are grown-ups, we can accept risks, we don't want to live in a nanny state. On the other hand, when we are pressed about individual risks, many of us expect the state to manage them, and again, when we are pressed, our tolerance of risk is usually extremely low. And we expect governments and 'those in charge' to manage those risks, although we don't trust government and those in charge, and we think we know best. So in short, we, the public, have dual attitudes to risk. I make the point because I think it is helpful, to try and understand public attitudes.

And we know that tolerance of delivery failure in national tests and exams is low. We saw that last year when there were errors on about a dozen questions over some 5000 or so exam papers. On the one hand, people could accept that mistakes happened, that a comma in a question might be in the wrong place. But they could not tolerate mistakes that made it impossible for a student to answer a question. Even if the mistake is just one mistake in a large and complex system, public tolerance is extremely low if a student is disadvantaged. There are other fields of endeavour where tolerance is low, and we must all recognise and accept that qualifications is one of those areas.

We are not unique – but we do have an unusual feature that heightens tension, and ups the stakes for those involved – results day. Patients don't all walk out of hospital on one and the same day. Results do come out together on one day, and that heightens tension, and public interest.

And when talking of public interest, it would be remiss of me not to mention the media. I recollect feeling dispirited once, on seeing an analysis of different newspapers' coverage of different sorts of public services. They varied in the accuracy of factual reporting, with some more accurate than others, but they were universally negative about public services, of whatever nature.

And in our field, it isn't always easy to get things across – what we do is complex, bafflingly so for most people, and when things happen it is more often as not the political, rather than the educational, correspondent reporting. Public attitude is influenced by what is read and what is reported, and we do of course recognise that.

So, we work in an area that is high volume, complex, high stakes and where media interest is high and risk tolerance is low. I'd like to look now at public confidence, and how things stand, starting with A levels.

We know from international and national research that Ofqual has done that A levels generally stack up well, and are generally well regarded. That doesn't mean that they are beyond improvement. In our research, people questioned the modular structure of A levels, and the resit opportunities that go with it, and we know from our

international research that assessment can get better, can be designed to better suit the subject, subject by subject. There are particular views about the subject content, in some subjects, and our research here was focused particularly on the views of those in higher education. We are looking now at responses to our recent consultation on A levels and we'll report shortly on changes we want to make, to improve them. But they already stack up well. And we know from our last, annual public confidence survey that confidence increased last year, and it remains high amongst teachers, parents and the general public.

So these are relatively stable qualifications, and stable specifications. We know where we stand.

For GCSEs, it is not so clear cut. These are the same concerns – about modularisation, and risks, and we are moving now, in England, to a linear approach. Controlled assessment, introduced to replace coursework, has not been trouble free, and we are revisiting that. That's not to say that we want to see a move to written examinations only. Rather we want the assessment methodology to match what is to be assessed, so far as is possible. And we want arrangements to be manageable for schools – not eating unduly into teaching and learning time. And we want the arrangements to be robust, to be able to withstand the pressure of accountability measures.

At Ofqual we've found some GCSEs wanting and we've taken steps to deal with that. This year's new GCSE geography qualifications are materially better than their predecessors, for example. And assessment in mathematics is better too.

But there is more work to be done - on the grading structure for GCSEs for example, where we question whether it is pitched right, and on the range of subjects available at GCSE, where we question where the line is to be drawn, if at all. We know that Government is consulting on English Baccalaureate Certificates to replace GCSEs in England in some subjects. We will continue with our work to strengthen GCSEs – they are important qualifications for young people and they will be around for a good while yet.

And what do the public think? Our survey shows a little less confidence than for A levels, and concerns amongst teachers and students about marking, I will come back to that.

But what about vocational qualifications? Ninety eight per cent of the awarding bodies we regulate produce 90 per cent of the qualifications we regulate – and they are vocational or professional or recreational qualifications. We know intuitively that public confidence in professional qualifications is high – although assessment experts in the room may question their validity because of some of the assessment methodologies employed. The picture is less clear cut on vocational qualifications,

and we need to do more to understand public perceptions, in sufficient detail, and more to drive standards – but of course, most of our time, our resource, is focused on academic qualifications, most especially when reform is in the air. That's quite a dilemma for us at Ofqual.

I'd like to turn now from public confidence to risks. We could talk all day, but I want to touch briefly on risks in three main areas; risks in reforming qualifications, risks in delivering qualifications and risks to standards.

Reform, first of all.

To state the obvious, reform is risky. And many of you here will have experienced it, I know, and played your part, first hand, in planning, designing and implementing reforms. There are lessons, from history, and we can look back at history to identify some of the risks, and think through how to best manage them. But one thing is worth saying straightaway: assessment experts and those with experience of reform will argue that it takes time, that fundamental reforms take six to eight years to deliver and 15 years to settle in. Politicians have much shorter timeframes in mind – always have, always will. We have to consider those different views, and find the right way forward. A way that does not put ministers' policy aims at risk of non-delivery, and more importantly perhaps, a way that does not put entire cohorts of young people at risk.

What does history tell us then about managing reform?

- Be clear about what you are trying to achieve – your policy aims – that's a first requirement, if you like.
- Engage widely, and build consensus, not just about the need for change but the nature of change and what will be involved for all players – most especially those in schools. We have something to learn here from the recent, successful recalibration of GCSE Science.
- Don't change things lightly, or for the sake of it – because continuity and stability are important, and because it is not always possible to foresee how changes will play out in the wider system. There are unknown unknowns, as they say.
- Plan, plan and plan. Plan carefully, and plan with others, across the system.
- Test - test as much as you can, as well as you can.
- Monitor and be prepared to pull the plug if necessary, if things don't go to plan, and sometimes they won't.
- Have a plan B.

I'd add one more thing – don't change too many things at once – because changes are inter-related, they affect each other, and managing change across interdependent systems concurrently is perhaps the biggest change challenge of all.

Moving on now to delivery risks in qualifications. Reform aside, it is here, in delivery, that many of the risks lie. We want delivery to be right, time after time, and every time. And we know that public tolerance here is low. We have a complex, distributed, large scale, interdependent system that needs to deliver examinations and results in an extremely condensed period. Mistakes will happen, but there will be little tolerance of them we know.

I hesitate to suggest it to this audience, especially those from exam boards, but regulation helps. We took action on exam errors last year, and things have improved. We took action when one exam board did not collate results properly, and things have improved. And we quietly take action to protect delivery of qualifications by other awarding bodies, those that are not exam boards, as well. I found myself authorising a dawn raid of an awarding body's premises recently, for example.

This is all fine and dandy, but it is simply a regulator reacting sensibly and as a regulator should, when delivery risks materialise. It is mostly reactive – reacting to issues as they emerge, or as we find them through our monitoring. And indeed it was said, in evidence to the Select Committee for Education earlier this year that Ofqual is a “crash scene investigator rather than an air traffic controller”. We understand that criticism. We do have to deal with accidents when they happen, but we also need to be on the front foot, we know.

And our regulatory framework and monitoring help. The way we work together to manage known delivery risks, for example coordinating the GCSE English November resit, helps. Exam boards know they are responsible for the safe delivery of qualifications, but we have a role to play, to ensure so far as possible that delivery is managed well.

The reform of the market with the introduction of franchising will put particular pressures on delivery, and will likely increase delivery risks. We need to work hard with exam boards to mitigate those risks.

I have in mind particularly here those qualifications that will not be in the immediate spotlight. GCSEs in some subjects for example, and A levels. I have in mind the cross-subsidies across qualifications in the existing model and the financial turbulence inherent in market reform. And the risks to the viability of some exam boards, depending on how things fall. All qualifications need to be delivered, and delivered safely, through reform.

Let me move now to standards, the third area I want to cover, in talking about risks to qualifications and to public confidence.

I hesitate to define standards for such an audience, but I'll start by saying that there's no 'one' standard, and no one way of defining standards. For ease, and for today,

let's look at standards in three ways: standards of outcomes, grades, results; standards of assessment, and content standards and the related issue of 'demand'.

Let's start with outcomes. Ofqual's view is that the concept known as grade inflation – the year on year increases in achievement that cannot be wholly explained – has done more than anything, over time, to reduce public confidence. And we have been addressing that starting on AS, then A level then GCSE. A level outcomes have stayed steady for two years now. It is a new world. And this year GCSE results steadied after more than two decades of year on year increases. We know, of course, of the concerns about variances in results for GCSE English, although outcomes stayed steady overall. We are working hard to get to the root of that, and we will report shortly. But I hope that those concerns do not detract from, do not prevent us from recognising, what has been achieved over the last two years, using the best methodologies available, to steady outcomes, year on year, for key qualifications.

Outcomes don't always stay steady, and shouldn't always stay steady. Sometimes they move, because the standard is intended to move, as in GCSE science this year. And we know that to manage a change in outcome standard – to raise standards – is a thoughtful business, and requires engagement with schools, and students, over time, as we did here.

Let's move now to standards of assessment, and touch also on the structure of qualifications. Modularity at GCSE has undermined confidence in the qualification, and we can see no standards reason why these qualifications should remain modular. Change is happening, as you know. Respondents to our consultation on A levels have reported similar views, in the main, although many favour retaining AS, for a number of understandable, and well-rehearsed reasons. We have more work to do there to collate responses and provide advice to the Secretary of State.

Most hearteningly, what we had to say about improving the quality of assessment at A level chimed with many respondents to our consultation and I'm really pleased about that.

Concerns have been reported by the HMC about the quality of marking, and indeed we are reviewing marking, and listening and acting on what HMC and others tell us, but there is an associated and equally important issue about the quality of assessment instruments. Marking will only be as good as the assessment being marked. You, more than many, understand the trade-offs here between validity and reliability, between predictability and greater challenge, between transparency and less transparent arrangements. We need to expose those trade-offs more, talk them through and so far as possible, gain consensus about how the balance is struck.

Some of you participated in our study of reliability. Thank you. And some of you are participating in our work now on validity. Thank you. These studies are helpful, illuminating, in some ways ground-breaking. But they will be of the most benefit if we can follow them through in the way qualifications and assessments are designed and evaluated. There's more to be done.

I turn now to the question of demand, and subject content – perhaps the most challenging aspect of standards in key qualifications today. For A levels, Government wishes higher education to be more involved in determining subject content, and most people responding to our consultation agree.

But the trick is to find ways in which higher education can be meaningfully involved, ways in which higher education agree, and ways in which exam boards can demonstrate that their A levels are designed with the right input from higher education. That's not straightforward, of course not. And again, there's more work to be done. But comfortingly, those in higher education responding to our consultation tell us that A levels are not off-field, in content. They are in the right ball park, and this is about keeping them in the right ball park for the future. Making the right adjustments to content, and the right trade-offs.

At GCSE, the situation is not so straightforward. The Select Committee for Education recently recommended that Ofqual hold the ring on subject content, by establishing subject committees. We are not empowered, or resourced to do that. QCDA was. We are not QCDA. But we can see the problem: responsibility for subject content is distributed at the moment. It is a rather organic model. And that may not be sustainable over time.

And government is now consulting on EBCs. And questions of subject content, demand, assessment and outcomes arise, of course, they do. And they need to be discussed, and resolved. These relate necessarily to the policy outcomes that Government seeks. We are discussing these things with Government, as you would expect. And the nature and extent of change will dictate the timeframe for delivery. Our job is to advise Government on the risks to delivery, and the risks to standards and to advise on a deliverable approach and timeframe.

And we need to get that right, otherwise public confidence will suffer.

Thank you very much for listening. Thank you.

