

CAMBRIDGE ASSESSMENT

SEMINAR

SCHOOL EXAMS: HAVE 'STANDARDS' REALLY FALLEN?

held at

RSA HOUSE, 8 JOHN ADAM STREET, LONDON

on

Thursday 29 April 2010

MR BENE'T STEINBERG, in the Chair

---

(From the Shorthand Notes of:  
W B GURNEY & SONS LLP  
Hope House  
45 Great Peter Street  
London SW1P 3LT)

CHAIRMAN: Good morning, ladies and gentlemen. A few housekeeping points first.

We are not expecting a fire so any alarm will be real. If you hear the alarm, please leave immediately via the exits which are there (Indicating) and immediately behind me and proceed to the assembly point in John Adam Street and Robert Street on the right-hand side of the Adelphi Buildings. Do not use the lifts and do not stop to collect your belongings.

Secondly, can you all please turn your phones, BlackBerries or PDAs to off or silent. Anything else can, and will, affect the PA system.

And finally, after all that, let me introduce myself. I am Bene't Steinberg, Group Director of Public Affairs at Cambridge Assessment and your chair for this debate.

I am delighted to welcome you all here on behalf of Cambridge Assessment. The group, which is a department of the University of Cambridge and a not-for-profit organisation, is made up of three exam boards - OCR in the UK, Cambridge International Examinations and Cambridge English for Speakers of Other Languages, together with the largest research capacity of its type in the world.

Today we will be hearing about school exams and debating on what is happening on standards today. Examination standards, and the perception of them, are of principal concern to society and dominate many educational and media debates.

Questions we expect to address today, but that rather depends upon you, include who owns the standards? Should standards rest on links between schools, higher education, employers and awarding bodies or be mediated through government and its agents? Just how much information is required to maintain public confidence?

Some ground rules. This is a debate, not a question and answer. If you want to pick up on a point made by someone else in the audience, feel free. If you want to make a statement, not ask a question, do so. All I ask is that you do it through the chair - me - and you make your points as concisely as this complex issue allows. I will cut people off if I think they have made their point already.

Secondly, you will note we are filming the discussion for the benefit of a much wider audience that could not make it to the RSA today. Please ignore the cameras but please make sure, if you are asked to speak, that you wait for a microphone to arrive in front of you, stand up (that is at the request of the cameras) and please give us your name and the name of your institution.

Let me outline the day for you. Each member of our distinguished panel will give ten minutes of their approach to the issue. We will then have a maximum of 15 minutes for clarification and questions. Then we will break for 20 minutes for coffee which will give you all a chance to decide when and where to weigh in on the debate. In your packs you have question sheets and those who wish to submit discussion points in that manner can hand them in at the beginning of the main session which will take us through to one o'clock.

Some interested parties have already supplied comments and questions via the web and I shall feed these in when I feel it appropriate. We are also streaming the event live over the

Cambridge Assessment website - and some comments may come to us via that route, which I will also add in when appropriate. I would encourage those of you watching at home, as it were, to submit your questions on-line.

This debate actually started on-line three months ago, beginning with a paper from our Group Director of Assessment, Research and Development, Tim Oates. Tim's paper can be found in your packs and Tim is going to start us off. Tim?

TIM OATES: Good morning everyone. Thanks Bene't.

You will be pleased to hear that we are going to clarify this whole area by making it far more complex! In essence, that is the challenge. We are concerned that the public debate is not sufficiently sophisticated to engage with the underlying issues which provide great challenge to us as an awarding body but also to the education community as a whole. Why the debate? Why should we raise this matter now in terms of standards within public examinations? It is essentially because of two issues. One is the fact that from various sources, in terms of the research community and the development community, there is evidence of the existence of mechanisms which could create grade drift in the operational processes of examinations, so the possible existence of those mechanisms is critical. I will outline some of them in just a moment.

The second is that there is evidence from diverse sources, not one of them definitive in terms of providing the simple answer to the simple question: are standards going up or down? So there is evidence from diverse sources, enough to stimulate anxiety, and this is what Simon Lebus raised in his response in the on-line discussion; that there is sufficient evidence to generate concern and therefore open up a public debate.

The first thing we want to do is to promote clarity, not by rendering the debate impenetrably complex but by clarifying some of the fundamental issues associated with standards. The first thing is to be very clear about which particular aspect of standards or what block of standards one is actually discussing.

We would separate the following: standards of demand, ie the demand which is represented by a particular examination and the questions within it, so what is being asked of students, and content standards. Of course we have to ensure that qualifications are up-to-date, and so content standards change over time as one necessarily has to update qualifications. However, that is to do with the stuff which is in the qualification and whether it actually meets the requirements of society, the economy and individuals. Thirdly, standards of attainment, the things which young people, candidates, actually know, understand and can do. We do need to be very clear when we discuss whether there is drift in standards, either up or down, in which respect are we actually discussing this very, very slippery concept of standards.

We then argue that one needs to be interested in standards for a variety of reasons and purposes. One is standards over time. Some panellists will argue that it is an irrational fixation to be concerned with holding standards static over time and what one needs in society and in the educational and training system are qualifications which are fit for purpose, so holding irrationally to an old-fashioned standard actually renders our system inflexible and unresponsive. Nonetheless, we do have some concerns about standards over time. After all, individuals apply to universities in the same year with qualification outcomes which have been derived by examinations taken in different years.

There are standards between specifications. An individual awarding body has different specifications in the same subject, at the same level, level two or level three, and indeed, different awarding bodies offer specifications, examinations in the same subject, at the same level. They should be comparable to ensure that fairness and justice is delivered across the system in terms of access to a particular grade.

Then of course we have the very thorny question of standards between subjects - should physics and mathematics, as cognate subjects, be of the same level of demand. Should psychology and physics be of the same level of demand and so on. We would challenge the notion that there should be a rational pursuit of some notion of common standards across the entire system in terms of subjects.

There are issues of standards between types of qualifications, academic and vocational, and of course there are issues of teaching standards and standards of education, but I am not going to deal with those today.

Briefly, what on earth are those mechanisms which I referred to earlier, the existence of mechanisms which could contribute to grade drift? They are many and various and this is where we need to become more complex in our thinking to understand what is actually occurring. We have had a period of constant change in the structure and content of qualifications, and one of my concluding points is that if you effect continual or inappropriate and unnecessary change of qualifications, it makes holding any standard extremely difficult. Maintaining standards in times of change is one of the most challenging things that an awarding body has to confront. We have elaborated the mechanisms for doing this but if you invoke unnecessary change in qualifications, either form or content, you are providing an unnecessary challenge to standards' maintenance.

We have all worked assiduously within the assessment community at removing bias, improving the quality of our exam items and improving the transparency of the assessments, removing what Denis Lawton used to refer to as "ambush assessment", hitting people hard with things that they do not expect at a time that they do not quite expect it.

The third issue is benefit of the doubt. The system works in terms of justice towards ensuring that benefit of the doubt is given. The work that was done by Desmond Nuttall, many years ago, and by Mike Cresswell actually looked at the extent to which there may be subtle processes within the awarding process where each year you can get a very small increase in the numbers attaining a grade based on the benefit of the doubt.

There has been vast improvement in the availability of materials - mark schemes, text books - driven by the performativity agenda. The concern for deriving a good grade has increased enormously as we have increased the pressure on schools and teachers in terms of accountability, and that has meant that there has been a dramatic rise in requests for clear materials about what it is that you need to do to attain a grade. That will lead to elevated improvement, just in terms of the efficiency of the processes around preparing people for examinations, but that does not necessarily represent an improvement in underlying attainment.

There has been a strong emphasis on inclusion across the system. Awarding bodies have delivered on the inclusion agenda. Again, this drive towards inclusion can increase the

numbers getting higher grades. Of course the whole system is, as Michael Barber in this very room five years ago suggested very strongly, oriented very, very tightly on examination performance in terms of making the system more accountable and more effective educationally.

Modularisation has an impact. It encourages boys, who might perhaps leave everything until the last moment in terms of an examination, to have to work right from the first few weeks of the course. They are attaining more, and a higher number of higher grades will be the result.

But also the routes through qualifications have become more complex. In many ways this means that young people can optimise their progress through education and thus optimise their attainment, again leading to an increase in the numbers attaining the highest grades. Crucially we have changing cohorts. Some of that change is just because people are choosing subjects in which they will excel or they get feedback from their AS qualification in the first year of their sixth-form studies which tells them that they are good at that subject and bad at that, even though they thought it was the other way round when they started the course, and therefore they can then optimise their performance when they move on to A2 in terms of subject choice.

There has been a massive investment in education. Undoubtedly, Michael Barber's increase in pressure in respect of the performance of the education system was followed by resources. The problem is that within that list of mechanisms there are mechanisms which just result in an improvement in the numbers attaining higher grades, without an improvement in the underlying standard of attainment of young people, and there are mechanisms which indeed do contribute to a genuine improvement in underlying attainment. The critical issue really is how much of each?

Just to conclude - where to from here, in terms of my perspective? We should understand with precision the impact of efforts to improve the standards of attainment, separating out the contribution of one type of mechanism from another type of mechanism, so we can genuinely understand how our efforts to improve the underlying attainment of young people, including the financial resources we put into that effort, is actually paying off.

Secondly, we have to reduce the frequency and scope of change in qualifications. That frequency and scope of change is driven by unnecessary change in respect of the content and form of the qualifications: arbitrary change is not helpful; frequent arbitrary change is extremely unhelpful in terms of maintaining standards. We are not arguing against change in qualifications - of course they have to be updated - but unnecessary change threatens standards.

As Bene't implied, we believe that qualifications should be owned through partnership between schools, higher education, employers and awarding bodies. Opening up this debate is not intended to say that awarding bodies know best in respect of all aspects of qualifications. What we are arguing is that awarding bodies have particular expertise in terms of measurement. In terms of the content of qualifications, their location within society and within the economy, and their role within selection, in order to deliver that, awarding bodies have to work in close partnership with subject communities and with the users of qualifications, but the notion of partnership there is absolutely critical.

What should drive it all? Fitness for purpose, clarity of purpose and validity. Those should be the things that we are concerned with when we discuss standards.

So what I would argue for strongly is we should resist illegitimate change in the format and content of qualifications; we should differentiate the contribution which is made by the different mechanisms we outline, and on that basis we hope that we can elucidate the factors and issues at play in order to better inform the public debate and the key question of whether standards have gone up or down. Thank you, Chair.

CHAIRMAN: Thank you, Tim. We are now going to hear from Professor Roger Murphy, who is an educational researcher with an international reputation for his work in the field of educational assessment and evaluation. He is currently Director for the Centre of Developing and Evaluating Lifelong Learning at the Institute of Research into Learning and Teaching in Higher Education at the School of Education at the University of Nottingham. Professor Murphy is also Director of the Visual Learning Lab, which is a HEFCE-funded centre for excellence in teaching and learning. He has worked in higher education for over 30 years and during that time he has been President of the British Educational Research Association, a member of the Educational Panel for the National Research Assessment Exercise and Dean of Education and Head of the School of Education at the University of Nottingham. He is the author of a large number of books and articles about educational assessment. Roger?

PROFESSOR ROGER MURPHY: Thank you. Can I start by congratulating Cambridge Assessment for having the courage to organise today and to let a few independent voices on to the top table. I think this is a very important debate we are having today. It is a debate that in many respects we have shied away from in the UK over a long period of time. I think we need to have the debate.

I have taken a slightly provocative take on the title of today and called my input “School exam standards certainly have not fallen”. That is addressing the simplistic notion, the kind of popular media notion that when the results come out each year, if there has been any kind of improvement, standards must have fallen or indeed if the results are slightly less good than they were the previous year, that also proves that standards have fallen. We do have some newspapers who are very keen on the idea of standards falling.

One of the things that needs to be said, and it is tricky to know who is on what side in this debate, is that in the UK we have a highly sophisticated system of public examinations. The irony is that people around the world tend to look to the UK for advice when they are setting up exam systems. People who have sophisticated exam systems of their own often look enviously at our system. What is at the heart of that? It is a multi-billion pound enterprise. It is very, very professional. One of the characteristics of our school exams in the UK that people are envious of is the relationship to the taught curriculum in schools. We are not in an American testing scenario where kids do tests that relate vaguely to what they do at school but not at all well. We have a tradition in public exams of trying to follow the curriculum, to update the curriculum and make it as relevant as we can, and have good-standing examinations that relate to that curriculum. That I think is both a strength and part of the problem because it is very complex when you are changing the curriculum and you are changing the assessments to match that curriculum.

The exams we have today - GCSE and A level - have a long history going back nearly 60 years. I know GCSE does not go back that long but A level does and GCSE relies a lot on the

prior history of O level and CSE exams. As I have said already, those examinations have not remained static over that time. If they had, the simplistic standards debate, as Tim has already said, would be much easier. If we just stuck with the 1952 papers, or whatever, and kept giving them year after year, then in one sense the comparison would be more straightforward, but it would be a ridiculous thing to do because the curriculum has changed so much and indeed we have got better at assessing students, so for a whole variety of good reasons the exams have had to keep changing.

The thing that needs to be said today, and in the lead-up to this I have been a bit misrepresented in some of the papers, is that in 2010 the examinations that students do this year will be a serious, rigorous hard test of the learning they have been doing in schools, and if the media want to attack that then I think it is unfortunate.

Then we come to the nub of perhaps where my position might be a bit more different from Tim's. What I want to say is that, despite all of those strengths, the UK examination grades are, and can ever only be, approximate indicators of student achievement. There are a whole load of reasons why that is.

One of the easy things about our exam system is we end up with very simple results. We try and reduce everything to a few letter grades. We try to summarise the complicated pattern of educational achievement of an individual child into a letter grade. Many of you in this audience today work very closely with the exam system and you will know jolly well that two students with a grade C in GCSE mathematics may have virtually nothing in common with each other at all. Even if they have done the same board and taken their exams through the same awarding body, they might have got 43 per cent by totally different routes; the mathematics they know is different. We have a simple grading system. We add up all of these grades. We look at comparisons between schools, between local education authorities. It is a very simplistic approach. But go back to the actual individual grade that an individual candidate gets; they have sat an examination - and often it will be a formal, externally set examination - which has sampled the syllabus that they have been following. Any examination is a sample and we all know that we can be lucky and unlucky in the sample that is chosen when we go into an exam room. Your heart sinks when you start reading the questions: "I've revised all the wrong things." This is one of the arbitrary things of examinations for the individual. They may be very lucky or they may be very unlucky with the sample of questions that they are presented with. Give that same candidate a different sample of questions on a different day and they will almost certainly get a different grade. It is even more complicated than that, and most of you know this too well. Give the same candidate the same examination paper on a different day and they will perform differently. We have good days and bad days, we are human beings; we are not the same every day of the week.

And of course it is more complicated than that because that takes us into the world of marking and grading. Some of my own research in examinations has been about the marking of examinations. We have a good body of work now showing that even with the most professional approach to marking that we have in our awarding bodies, you cannot in subjects, particularly the ones that rely on essay-type answers, have a system that is absolutely accurate and reliable. Markers are asked to use their judgment. They are given advice and standardisation but they vary and all the published work shows that that variance is quite considerable and means that we have to be careful in the use of an exam grade because it is

not an exact precise judgment; it is an approximation. You could easily have got a grade higher or a grade lower for a whole variety of reasons if you are an examination candidate.

When we try to make comparisons between subjects things get even more complicated. We give the same letter grades in different subjects at A level/GCSE. We say that a grade A in mathematics is in some way like a grade A in chemistry and in some way like a grade A in French. That just cannot be the case; they are different subjects. It is a chalk and cheese discussion. If you are comparing a grade A in French with a grade A in mathematics, these are two completely different things. It is probably rather misleading to call both of them grade A, but we do it. We cannot have precision in that area.

The same thing really applies to the thing that Tim and I agree about is one of the most difficult aspects of this debate about examination standards which is the standards over time debate. I do not think our public examination system is the best way to address standards over time. That is not its prime purpose. Its prime purpose, I would argue, is to give grades to individuals to help sort out what they are going to do next: at GCSE, whether they go on to A level; at A level, for many of them, what kind of employment or higher education they might enter. This idea of looking at patterns of grades over as much as 60 years and trying to sort out whether education in England and Wales and Northern Ireland has got better on the basis of this year's results; the results will never answer that question. Okay - have a look at them, but there is so much complication behind them: the number of children staying at school now, we are delighted, has gone up; the curriculum has changed out of all recognition; people have got smarter about preparing for exams; the awarding bodies have got better about allowing candidates a range of ways of showing what they can do. The thing has changed out of all recognition so these crude comparisons over years mean nothing at all in terms of the detailed standards over time debate.

My final point is to summarise and say that we need to treat exam results as approximate measures. The problem with this debate is people try to read too much into exam results and use them to try and answer questions that they cannot answer. Thank you very much.

CHAIRMAN: Thank you, Roger. Moving from somebody very heavily into the research field to perhaps a bit more of a practitioner: John Bangs, Assistant Secretary of Education at the National Union of Teachers. John studied Fine Art at Reading University under the tuition of Claude Rogers and Terry Frost RA. After gaining his degree, he completed a teaching course at Goldsmith's College and spent a number of years teaching art and literacy at Bow Secondary School and Templars Specialist School, so he has an enormous amount of experience on the ground. John?

JOHN BANGS: Thank you. I am very, very aware that there are colleagues from other teacher organisations in this room looking at me in a kind of glinty-eyed way! I am going to have to be very careful, Chairman! Also I want to apologise because I have to go at quarter past 11 because Vernon and Ed have given us a job to do, which is to tell school governing bodies "to instruct head teachers to remain absent from school at times when the tests are due to take place while another person administers the key stage two tests". That came out last night and we are going to have to consider the implications of that. The effect on our NAHT colleagues in particular I can only guess, but it probably will unify governing bodies and head teachers in a way that nothing else can! So all good for that.

And I would also say thank you very much to Cambridge Assessment. We have a long-standing association. Tim came to speak to our National Education conference last summer in the graveyard slot on Saturday afternoon and no-one fell asleep, which was fantastic actually, it was great, it was a really super performance, and the debate that Cambridge Assessment has raised is absolutely spot-on.

Your key question, ie, what do we mean by standards? is absolutely right. I had a colleague who was quite rigorous and I remain in his shadow on this. We should be talking about performance against standards all the time, but “standards” is a word that has been used inaccurately by politicians over the years and conflated with the idea of performance. I agree with Roger as well that a discussion about how you set the standards is separate from examinations.

To move on then and some reflections. The quality of what goes on in our schools is the best that it has ever been; unequivocally we believe that. Of course it could be better, of course pockets could improve, of course the debate about those not in education, employment and training is absolutely right because it is addressing a very real problem, but overall the quality of what has gone on in schools in terms of pedagogy, in terms of knowledge, skills and teaching is the best. It could always get better. I hope it is not as good as it gets in the light of what could happen after the General Election in terms of cuts, but it is the best and for that we rely on Ofsted. We might have criticisms of Ofsted in terms of the high-stakes consequences of Ofsted inspections, but what else can you rely on in terms of the number of schools who have a notice to improve being reduced. The big question then is whether the examinations we have got are appropriate for measuring that improvement.

I think we need to have a look at the big picture really because there are countries, including Sweden, that do not have a highly complex examination system, not in any way, shape or form. In fact, they have an end of phase evaluation which is internal. You could say that is good or you could say that is bad, but they are not hung up. I say Sweden because that is the favourite *de nos jours* for edu-tourism for all of our politicians at the moment. The debate we are having about the nature of qualifications is not particularly an international one and countries are peculiarly focused, particularly in terms of upper secondary achievement, on their own examinations and fixed, and peculiarly it is not an international debate about the nature of qualifications, it is very, very much focused on what goes on in individual countries.

To move on, there is that question about are the qualifications appropriate. I will say this because I know other people will say it as well but it is worth us saying it, I did think Tomlinson came the nearest to asking about the nature, relevance and purpose of qualifications. However, it is a big political hot potato and I remember Barry McGaw’s study for QCA on standards over time where he basically vindicated the argument on standards, ie the setting of standards or performance against standards had not gone down, but there are other bits and pieces of studies which have shown that there have been some question marks over individual subjects, such as science, for example, which I think have been sorted. There is that bit there.

Then there is the next bit, which is the involvement of the teaching profession. We are the only teacher organisation that puts money where our mouths are in terms of the standing joint committees. I expect quite a lot of you know about the standing joint committees, but they do relate to the awarding bodies and they have creaked along over the years and some awarding bodies - not OCR and I will not mention the other one - have not been at various times

particularly enthusiastic about hearing what teachers have to say about examination papers over the years, although that has steadied out. I have to say that neither have some teacher organisations been that enthusiastic about capturing the professional views of their members about the quality of something so fundamental, not only to young people's achievement but also to what they teach in the classroom in secondary schools.

I do think there is a major issue here and that is that the reality is in terms of looking at what is in the specifications, what is in the examination papers over the years, the profession is not integrated as much as it could be in terms of a commentary system about the nature of specifications and examination papers themselves. The fact that we have to put in a paltry 2,000 just to keep the creaking edifice going is a sign that that is not the case, so we do need to have a look at that.

Also, I think there is something else we do need to say and that is that there is a very, very deep yearning for continuity. This is where we have a little look at the manifestos. There is a line in the Conservative Party's manifesto which says examinations should be more robust and we need to give universities and academics more say on their form and content. There is a very big thing about an implication about the Bac and IGCSE, their qualifications, about which they are very enthusiastic. The words 'more robust' indicate the Conservative Party's belief and concern by implication that somehow standards or the standard of the qualification has gone down. The Labour Party's manifesto talks about the review we all know about in 2013 in terms of the whole qualification system including diplomas. The Liberal Democrats have come up with a very strong view about having the Tomlinson general diploma and an Educational Standards Authority.

Some of these have merit and all of these have a political implication to them, but I have to say there is a deep yearning for continuity amongst the profession, amongst schools. Also, I have to say - and I do not feel I have to say it actually - we need to ensure that Ofqual works, we really do. The relationship between Ofqual and the teacher organisations is a good one and a developing one. We do not have a problem with access to Ofqual. If we want anything we will talk to Ofqual. There is a very deep yearning that the politicisation implied by the manifestos is something that we really, really do not need.

There is a debate about the nature, as Roger said, of the standards set and the nature of the qualification and what is the best way it seems to us. There is a strong argument for looking at why, for instance, the Liberal Democrats propose the Educational Standards Authority. I think there is an implication there that standards and what goes on in schools in terms of teaching and learning really do need to be more integrated, for example. There is an argument for looking at GCSEs as Tomlinson did, but the idea that somehow we have to return to a tiered system where children are automatically sorted at 14 between vocation and academic routes is the last thing the vast majority of colleagues in the profession want. Thank you.

CHAIRMAN: Thank you, John. Switching back to science, as it were, Gordon Stobart is Emeritus Professor of Education at the University of London Institute of Education and Professor of Education at the University of Bristol. His research has focused on assessment, national systems and policies and classroom-based assessment, especially Assessment for Learning. He has been involved in several assessment-related projects funded by the Nuffield Foundation and the UK Department for Education and Skills. He is a founder member of the United Kingdom Assessment Reform Group and his most recent book is *'Testing Times - the*

*uses and abuses of assessment*', which seeks to reclaim assessment as a constructive activity which can encourage deeper learning. Gordon?

PROFESSOR GORDON STOBART: Thank you very much. This is a bit like one of those plenaries where you report back from small groups and if you are the person who is reporting back further down the line, you cross off everything you had on your list and then when you have crossed off everything, you think the people who have gone before have said everything I wanted to, but I am going to say it anyone and that is where I find myself. I want to emulate the politicians here and I want to say, let me be quite clear about this, the answer is yes and no to the question have standards really fallen. If you say yes, I can find you, I am quite sure, a 1970-odd A level paper that has harder questions than we will find in an A level paper today, not difficult. I am thinking particularly of 1975, a vintage year, when 44 per cent of students - and the students only represented about 15 to 20 per cent of the school population - failed maths. That is what I call rigour and that is what I call standards! It was pointed out the reason so many failed was that they were not mathematicians, they were people wanting to do geography, economics and other things who took a maths paper which was completely inappropriate and therefore failed.

I can also find O level science questions that were about real science, which is about memorising formulae and being able to use them in an exam. Yes, if we look at those sorts of things we can say things have got easier. The 'no', and I think it is a big no, is that if we are talking about the percentage of 16-year olds and 18-year olds, I would argue that the underlying knowledge of a large percentage of these has increased over the 20/30 years. If we are talking about the 14 per cent going into higher education in 1975, we have now got something like 45 per cent going on to higher education. If we take the average student in 1975, somebody at the middle of the attainment distribution, they finished school at 15 or 16; now they may be staying on to go to A level. This leaves me with a Zen-like thought for you: that sometimes you have to lower standards in order to raise them. If your population is changing, there is an access agenda, you may have to do things like this to engage a larger proportion. I think you can have both, you can have easier questions and standards-raising. That is my Zen moment.

My favourite analogy to try and capture what Tim and Roger were saying about this is climbing Mt Everest. In 1953 two people, we know them well, got to the top of Mt Everest. There was a day in 1996 when 39 people were on the top of Everest at the same time - 39 people on the same day. What has happened to Everest over time? This is my question! I have engaged in a debate because I asked has Everest shrunk and you will be pleased to hear it has got bigger, it is 26 feet higher than it was 30 years ago. What has happened to mountaineering? If that many people can get to the top, surely mountaineering standards have dropped? What they have got now is things like maps, training, equipment, guides, all those sorts of things, so clearly standards have fallen. This analogy has run and run. Mike Cresswell did an analysis of the proportion per year who successfully got to the top - this is all about rigour - and it was really sloppy in about 1995 that something like 0.9 were getting to the top. In the 1970s when there was rigour they kept falling off and just not getting there.

What this risks is the idea of a pointless "golden age" debate about standards; when was the golden age of education? What tends to happen is whatever age is announced as the golden age, somebody at that point will be saying: "It was much better 30 years ago, much harder when I was in school", so there is no fixed point here.

A couple of quotes here and this may seem tedious, but I am going to read a complete examiner's report from 1924 in pure and applied maths: "The only point that calls for a report is the general weakness of a large proportion of the candidates." That is it! They do not mince their words up at the JMB. Things had not got much better by 1932: "A considerable percentage of the candidates were quite unfitted to take the exam and had no possible chance of passing." Another informative bit of feedback. One final one here: "It has been said, for instance, that accuracy in the manipulation of figures does not reach the standard which was reached 20 years ago. Some employers express surprise and concern at the inability of young persons to perform simple numerical operations involved in business." 1876.

I think my point for today is that the standards over time debate goes nowhere and we ought to drop it as soon as possible. What it should do is lead in to the proper debate, it seems to me, which is about the fitness for purpose of our current and future qualifications. The dilemma with this is we then hit the notion of multiple purposes. When you have got qualifications which are being used for several, in fact dozens - Paul Newton is here who is the purpose spotter and has worked up for 18/14 for the National Curriculum test - if we are talking A level, talking GCSE, we have got accountability, we have got driving up standards in schools and we have got individual selection. That is a problem in itself. When we are trying to operate with multiple purposes, an exam cannot do everything well. Alistair Pollitt uses the nice analogy of the Swiss army penknife - it is a great one this - it does everything but it does nothing well. If you have ever seen the saw blade on a Swiss army penknife, imagine yourself with a log on your leg and, "Oh dear! Oh good, I've got my Swiss army penknife", you would be found dead from exhaustion, not from an injured leg!

The notion that we can take a qualification and use it for so many purposes - are national standards rising and falling, are we driving up standards in school - again, we have got to get back to what is the principal purpose of this, and I think the suggestion today is that we are looking very much at the selective purpose of something like A level. Is it fit for purpose for selection? I think there is a crisis here in the sense of Tim's grade inflation and ratcheting up and things where we have got more students with three As, so that selectors cannot discriminate in terms of grades because more than enough have got the required or the very top grade. We need to do something about that because the exam in that sense is not fit for purpose if it cannot do that selective work.

It would be interesting to see whether introducing the A\* helps with this or other things happen as a consequence of this. I also note, and this could cost me my lunch, the Pre-U from Cambridge feels to me like back to the future. Are we not back to ten per cent taking a much more difficult exam and things, which is where A levels started out, but I will leave that there. Where the debate should be going is to ask questions about are we relying too much - and this is Roger's point - on grades that are not necessarily that robust or we cannot put our complete confidence in them. As the prime determiner of whether somebody gets to university, is this the right way around or should we be having a bigger debate about when you apply for university - I know there are UCAS forms and things like that - what else should be in the mix for selection rather than simply relying on exams, which I do not think can bear the weight of some of the purposes which are being put on them. Thank you.

CHAIRMAN: Thank you, Gordon. Our final speaker is Anastasia de Waal, who is the Deputy Director of the think-tank Civitas and Director of its Family and Education section. Her work centres of analysing social and economic trends and evaluating family and education policy, as well as formulating and directing Civitas' overall research agenda.

Anastasia is a qualified primary school teacher, specifically trained for inner-city teaching. A regular contributor to broadcast and print media, her publications include: *Inspection, Inspection, Inspection, Second Thoughts on the Family* and *Inspecting the Inspectorate*. She is the Chair of Parentline Plus, a board member of Women's Parliamentary Radio and a panellist for the *Observer*. Anastasia?

ANASTASIA DE WAAL: Thank you. As a few people have said, I think that Cambridge Assessment is having this debate at all is not just an indictment of its bravery, but also an indictment of the fact that this has become a debate which is not just about haranguing critics, it is not just about looking to the golden past; it is a debate which needs to be had on a wider scale. What I am going to do is not look at exam standards, I do not think exams are the problem at the moment, if there is a problem, but the way that exams, and particularly the pursuit of high grades in exams, is affecting standards of learning. That is what I think really is the core issue, the fact that a pursuit of A grades at A level, of reaching the desired standard in the Key Stage 2 SATS, of reaching a higher number of As at GCSE, the impact that that is having by the impact from the centre in terms of government input on the distortions in school. What I want to do is do a bit of a tour through some of the work we have done at Civitas looking at the different ways in which the pursuit of higher grades and effectively government intervention have led to distortions within learning.

To start at primary school, at the moment, and for quite a number of months now, even years, we have been talking about the problems with the Key Stage 2 SATS. The answer now seems to be to scrap the Key Stage 2 SATS, but I wonder whether that would solve the problem because ultimately the problem is about a huge pressure, from the centre again, for teachers to come up with the necessary results. Getting rid of the SATS, getting rid of testing I am not sure would get rid of the problem because there still would be pressure for schools to be able to demonstrate that the necessary level was being achieved.

In order to try and understand a bit about how teachers felt when it came to standards and the SATS, we did some research of Year 7 teachers, so looking at the teachers who get Key Stage 2 pupils into their classes and asking them whether they felt the results that the primary leavers were coming in with were reliable. 90 per cent of Year 7 teachers in our survey felt that they were not reliable and nearly 80 per cent felt that up to a third of their pupils in Year 7 were lower than their Key Stage 2 results. Whether that matters or not is really not the issue. Does it matter that it looks like pupils are doing better than they are? No, not really. Is it all right if government says that educational policy is having a better impact than perhaps it is? No, it does not particularly matter. What matters is the behind the scenes impact, and this is the recurring theme throughout the education system. The problem at the moment with SATS, if there is a distortion in terms of performance, is what is happening to the curriculum. The problem is ultimately that the curriculum is being shrunk in many cases in order that that snapshot of learning is becoming a sum of learning so that the necessary levels can be achieved.

Then moving into GCSE, that is another area where we have looked at. Here it is an interesting case because it is obviously a benign intention which in many ways has become part of this trying to chase higher grades and trying, in effect, to prove that education policy is having the desired effect, but through grades.

The equivalence system has been talked about a lot in the debate around standards, and whilst it seems like a very good idea and continues to be a good idea to broaden the curriculum in

terms of opportunities, one of the big problems is the way this broadening of the curriculum has led to a way in which the system can push weaker pupils into achieving higher grades but not by maximising learning. One of the pieces of research that we did fairly recently was looking at academies. Academies are held up as the beacons of success in terms of closing the achievement gap, improving life chances, particularly in inner city areas, and really making sure that learning is maximised. What we wanted to do was to find out from academies who are achieving a much higher rate of high performance at GCSE, what subjects they were doing. One of the anomalies seemed to be that academies are at the forefront of the political agenda, yet there seemed to be very little information about how this success was being achieved. The results were very revealing. The majority of the principals of academies who we interviewed thought that they did not want to release their results. Some of the reasons they gave for not wanting to release their results were because they did not want to expose the fact that they were focusing on some subjects and not other subjects. When they were asked whether they thought academies should have to publish a breakdown of their results, as other maintained schools have to, only 55 per cent thought that should be the case. Again, a lot of the arguments against publishing their results or a breakdown of what was being achieved at GCSE was because they did not want to show what subjects were being used to attain higher grades.

That in itself was not necessarily significant. What was significant was when we did look at the results from those academies which did give them to us. There were a lot of ICT, a lot of BTEC IT qualifications, exactly the qualifications which Ofsted has said are not rigorous enough, are not stimulating enough and are not wide enough in terms of their learning remit. In many cases there were very low entries, in some cases no entries at all, in history, in geography. Single sciences were very rare in the results that we saw and the important thing to remember is these were the results that we were given, so we do not know what the results of those academies, when broken down, were for those principals who did not show us their results.

Why that matters particularly is because we know that academies are serving the least affluent areas. If this then means that in the pursuit of higher grades there is a lessening of learning, then that is the opposite of widening opportunity and of broadening life chances. It might look good on paper but it does not look good for opportunities. It has been said that the important thing about equivalence is that it is great for the school, although in more cases it is not because the teachers find it very unsatisfactory, but ultimately the student does not walk away with an equivalence, they walk away with what they actually have, the breakdown of their results and perhaps a narrowed remit in terms of learning.

Again, last year we looked at A levels. A levels seem to be at the centre of the debate around standards, which is interesting considering that perhaps we should be thinking about standards in schooling where everybody is participating rather than at A level. What we were trying to understand - again, talking to teachers - was whether they felt that the A level was comparable now with the past and whether this idea of continuity in terms of comparison was valid. One thing was very clear - it is not and it is not that interesting to compare standards in that sense because we are talking about a fundamentally different system; we are talking about a modular compared with a linear A level. I think it is a very valid point, and a lot of teachers made this, in saying that we cannot compare standards because we are talking about a completely different qualification. That is very valid but it is not valid for government to say that things have dramatically improved if it is, indeed, incomparable. That was one important point.

The most important point was the impact of the modular A level in terms of re-sitting. That was the thing that came up the most in terms of why the modular A level was considered to be more accessible. The majority of teachers supported this idea of a more accessible A level. Why? Because they thought it was very important that more kids had opportunities to go into higher education. However, what a lot of them felt was a problem was it was re-sitting *ad nauseum*. It was re-sitting the same learning over and over again rather than being able to broaden what they were learning. Maybe it meant higher grades but perhaps it meant an impoverished learning experience.

It is also very important because some of them felt that although the idea was that re-sitting meant more opportunities perhaps to go to the university, universities were then turning around and saying, “Actually, we don’t think these results are an indicator which will suffice, therefore we are going to do our own tests”, so in other words the ultimate slander. The Government talks a lot, and I think rightly in some ways, about the insult to students every year when there is this debate about standards, but unfortunately there is no greater insult than a university turning around, legitimately because they have to, and saying, “Well, we can’t use your results because they are not a terribly satisfactory indicator.” What is also very significant is that 69 per cent of those A level teachers whom we interviewed said that they thought that more than 50 per cent of their cohort for that year had improved their overall grade in their subject at A level by at least one grade. In other words, it is having a very dramatic effect. That is the clearest thing when it comes to A level, that we need to be honest about the fact that we are not talking about like with like. Does that matter? No, not necessarily, it could actually be much better and this idea of more accessibility in A level is very important, but it does mean that it is not legitimate to say that things have dramatically improved.

Finally, the work I am looking at at the moment is higher education and I want to bring that in because this idea of perhaps the less intrinsic learning experience falling by the wayside and the emphasis being on the qualification or the grade itself is very significant in higher education. There is clearly an important agenda to get more people participating in higher education, but it has been described as box piling to an extent. That is that we are talking not about the intrinsic value of the learning experience, the intrinsic value of the degree that you are doing, but the fact that you have got this qualification. It does not matter what it is in, it does not matter whether you find it stimulating, it does not matter whether the quality is good, in other words it does not matter what is in the box, but it matters that we are stocking and stacking those boxes up and that is problematic.

That comes to the conclusion of what the situation in terms of a focus on grades rather than the learning experience is having in terms of the wider education situation. I think the greatest thing we need to do is sever national accountability on a policy level from individual student performance. The second thing we need to do is we need to make sure we are looking at standards from a wider perspective so that it is not just about exam results. That means particularly Ofsted looking at not just data in offices outside of schools, but the greater holistic experience of schools. Most importantly we need to make sure that exams and tests at all levels in education are not testing the sum of learning but genuinely the snapshot of learning. That does mean much less interference from the government, it means making sure that in primary school teachers are not told everything, they do not want to know what is in a SATS test, but they are being told because it helps to boost performance, and it means making

sure that A levels and GCSEs are about stimulating students and having a wider learning experience, not simply about exam preparation. Thank you.

CHAIRMAN: Thank you, Anastasia. Right, ladies and gentlemen, I think we can genuinely say that we have covered practically the entire waterfront of what we could debate after coffee. We have five minutes for points of clarification. The debating will take place after coffee. The discussion will take place then or the point-scoring, if that is how you feel you want to play it, will take place then. If anybody has a point of clarification they want to ask anybody up on the platform where they did not quite understand precisely what was being said, now is the time to do it. The same rules apply: could you stand up and give your name and institution.

TINA ISAACS: I am Tina Isaacs and I am at the Institute of Education. I have a question of clarification for Tim actually. At the very beginning you talked about, two or three times I noted it down, the “unnecessary” change to qualifications, and what I wondered was if you could give us a concrete example because, if you think about it, the new A levels that were introduced in 2008 were followed on from qualifications that were introduced in 2000. The new GCSEs that were introduced in 2009 followed on from qualifications that were introduced in 1999. So what I am wondering is how you define “unnecessary”?

TIM OATES: Thank you, Tina, that is a key question. One could look to particular instances of change which we have highlighted before in terms of some of our commentaries. For example, we know that in GCSE mathematics that calculators have been in/out, in/out a number of times and each of those changes then gives rise to problems in terms of maintenance of standards and what is actually going on within the qualification in terms of attainment and the inferences that one can make from it. I would argue strongly - and others would disagree - that changes should be led by the changes which emerge from the subject community or from the structure of knowledge within the discipline itself. That would suggest that we need to change subjects on a time-frame which is determined by the subject itself, as opposed to effecting total system change, so for example moving the system wholesale from linear to modular. Others would argue that it is easier to maintain standards when you change everything at once, and I think that is a critical point for this debate. In terms of change, what should be the main drivers? I would argue strongly ---

CHAIRMAN: Hang on, Tim, you are now getting into the discussion. If we could leave that. You have answered the point and you are now into discussion. This is just for questions. Yes?

GREG BROOKS: Thank you, Greg Brooks, Emeritus, University of Sheffield. I also have a question for Tim. Tim, the slide you had headed “Where to from here?” the third bullet was about who owns the exams and it did not mention government; why not?

TIM OATES: I think that one of the defining features of the system two decades ago was that the major transactions were between schools and awarding bodies and awarding bodies, higher education and employment, and there were frequent instances of where schools or groups of schools or LEAs got together to produce interesting and invigorated curricula which met with precision the needs of young people, and those clusters of schools or schools would then turn to an awarding body to say, “Can you help us with assessing this?” Really I consider that to be the right sort of relationship in respect of the determination of the content

and the format of qualifications. It is those kinds of transactions which should drive and transform the system.

CHAIRMAN: On that note, I am going to call for coffee, which I think will be valued by everyone. I will see you all back here in 20 minutes when the debate will kick off properly. If you have any questions, please feed them into the team or to myself, or comments, or anything like that, and I can take them from the chair here. Thank you very much, ladies and gentlemen.

After a short break

CHAIRMAN: Thank you for coming back, ladies and gentlemen. It is always nice to know that we have not lost far too much of the audience because it was getting tedious! I think this then becomes perhaps the most interesting part of the debate. I cannot resist starting with my favourite quote.

“It has now become clear that public opinion in England was disposed to put quite an excessive reliance upon the system of competitive examination as a panacea for educational delinquencies or defects. Examinations as ends in themselves have occupied too much of the thoughts of parents and teachers. Their very convenience and success led to their undue multiplication until they were occupying too large a place in the system of national education.”

From the 1911 report of the Consultative Committee on Examinations in Secondary Schools!

So this is by no mean a very new debate, but we are perhaps a tad closer to finding some of the answers. We now have a fair number of questions, some of which I will read out and some of which I will actually call the person to speak to. You will be pleased to know that we have over 700 people watching us on-line, many of whom are already sending in questions and comments as well.

I thought we might start with an on-line point just to throw that open to everybody. Questions related to: “Do we agree that the control in relation to standards cannot effectively be assured to be unbiased and impartial when provided for by a transient government or their bodies, lacking educational expertise?” We have had a number of those sorts of questions. Would anybody like to pick up on the interesting question of what role the state has in all of this, whether they are driving up standards or pretending to drive up standards, or whether they should be engaged in this matter in any way whatsoever? Many people would say that it is the role of the state to set goals, but how we achieve them is up to us. Does anybody want to pick that up or do you want to start the conversation completely differently? Yes?

AMANDA SPIELMAN: Amanda Spielman from ARK Schools. As you know, I just contributed to the Sykes Review report on this very subject so I would like to bring up what we said, which is we think the state has a very limited role in relation to the core subjects where we think there is a need to create a clear consensus on minimum standards, not necessarily on what constitutes every grade of achievement but what constitutes minimum standards for the main purposes. We would not see that extending out across all the range of subjects or disciplines that can be assessed.

CHAIRMAN: Amanda, when you say standards, as Tim started off, are you talking about content standards at this point?

AMANDA SPIELMAN: Yes.

CHAIRMAN: So the state has a role in setting minimum content standards?

AMANDA SPIELMAN: Yes.

CHAIRMAN: Okay. Would anybody else like to pick that up?

DAVID HOWE: David Howe, Enterprise UK. I am not entirely sure that I agree with that. I am not even sure they have a role there. Surely their job is to facilitate a professional dialogue between the education world and the employer world in terms of meeting needs and to ensure that is actually happening and is going somewhere. The fact that a government is in a position to even define minimum standards here, I am not sure is actually relevant because I would argue that many politicians sometimes simplify this very debate we are having today down to one-line comments for the *Evening News* or the tabloid front pages the next day. The level of debate that we are engaged in today, and the complexity, is something which is actually beyond, as far as I can see, most politicians. I would argue that I do not even know if government actually has a role in defining those standards. I think they have much more of a role of ensuring that there is debate and there is a meeting of minds, as it were, and meetings going on between various bodies that are interested in this and ensuring that it actually comes out with something.

CHAIRMAN: Anybody else? I am reminded that in the 1850s the Government said they did not need to supply an examiner to the very first board because they thought we could do it ourselves and did not need state intervention.

ANDREW HARLAND: Andrew Harland, Examination Officers' Association. As you know, there are three parts to the education system: the teaching of learning, the delivery of public exams and the assessment part. We feel that the debate that affects the delivery of exams in the country is not about standards but about the expectations that are placed upon students and the expectations that are placed upon the exam system. Certainly looking around the room at educationalists, awarding bodies and so on, really it comes back to Gordon's comment right at the very beginning as to whether the system is fit for purpose. That is really a focus in this debate, not just on standards but the expectations of what that system is supposed to deliver.

ISABEL NISBET: I am Isabel Nisbet from Ofqual. We are a little limited in what we are able to say today, but I just wanted to say that there is a regulatory role for the state which is different from the ministerial government role and (including in the Sykes Report) there are various views about how that could be exercised. There is a common theme about assuring the currency of the assessments or the qualifications or the tests, and the debate is whether the degree to which some arm of the state other than government can do that. The alternative could be a very mixed market. Historically, students have had to guess among a very varied set of offerings without any reliability of the currency with which these were being measured. There is a regulatory debate around this. We are listening very hard to what people are saying today and we will be talking to many of you after 6 May.

SIMON LEBUS: Simon Lebus, Cambridge Assessment. I would just like briefly to respond to what Isabel said about the role of the regulator. I think there is a very important regulatory role in terms of a weights and measures type function, but one of the difficulties, and one of the reasons we are having this sort of discussion is although I think we have in a sense got on top of the measurement science, we have not got on top of what is being measured. One of the reasons for that comes back to some of the comments that were being made earlier about the divorce between awarding bodies, schools and HE. The reason that divorce has occurred is that there was a large regulatory and political apparatus that has inserted itself in the relationship and, as it were, disintermediated those groups, and I think quite a lot of the challenge is in making sure that Ofqual or the regulatory arrangements are focused on weights and measures and that some of this emphasis on curriculum and curriculum development is as a result of an unmediated discussion between HE, awarding bodies and schools.

CHAIRMAN: Which interestingly takes us to Ian Glossup, who is watching on-line, where he lumps everybody together. "Is not the problem not the lack of public understanding - they understand just fine, thank you - but the failure of the awarding bodies and their political masters [in which I also include regulatory] to justify their claims with objective data?" which I think goes to the heart of the entire prospect because the gentleman back there was talking about instant media responses by political masters. Is the fault that we are not actually giving the public and/or the users of exams - HE - enough data for the purposes of using them?

GREG BROOKS: Greg Brooks, Emeritus, University of Sheffield. In the interest of introducing some data into the discussion I wonder if I might indulge myself for about two minutes in summarising the work that my Sheffield colleague Sammy Rashid and I have done. We have collected together and analysed all the data on the literacy and numeracy levels of attainment of 13 to 19-year-olds in England from the very first national survey of the reading levels of 15-year-olds in 1948 up to last year. It is quite a long story so I will not try to read it all out to you. The short story is that there appears to have been a rise in the average level of reading attainment for 15-year-olds between 1948 and 1960 and then in both reading and then writing, when it came on the international surveys, and numeracy, a long plateau when average levels of attainment barely shifted other than in very small details. In about the last 15 years there have been steady and, in some areas, substantial rises in the percentages of young people reaching particular levels or above in vocabulary and maths at age 14, in key stage three English and maths, and in GCSE English and maths, and those are substantial achievements in the system. They mean that, amongst other things, the top levels of attainment of young people in our country are amongst the best in the world. At the other end of the scale there is, we think, a problem; namely that about 17 per cent of young people have less than functional literacy (I am talking here about 16-to-19 year-olds) and about 22 per cent of them have less than functional numeracy, and that really is a problem for the system and there need to be measures to reduce those tails of under-achievement. Thank you.

CHAIRMAN: Roger, you wanted to come in on that?

PROFESSOR ROGER MURPHY: If I may. This is a familiar bit of the debate where people say, "Oh goodness, can't we just have a simple answer to this question? Can't the awarding bodies give us the data to see whether exams have got easier or not?" It is a fair enough question to ask but the problem is more complicated than that. There is no shortage of data. There is a huge wealth of data, like the stuff that Greg is talking about. The awarding bodies themselves have done huge amounts of research into comparability issues. The problem is in the question, as we have said already today, that the question is never as simple as it seems. If

people are interested in standards over time they are asking about a very complicated set of issues. If you want to compare A level standards in 1950, when they first came in, to standards now, so much has changed, as we have been saying this morning, that it is not a simple question. It is a very complicated question and so it is no good saying to the awarding bodies try a bit harder and come up with the data to answer the question. The question is too complex to have a simple answer.

ANASTASIA DE WAAL: Just one thing on that point. People are interested because it has become a sort of “pop quiz” question whether this year’s A level is comparable to ten years ago but I think people really are interested in how educated people are. The bottom line is are people better educated, are education levels the same, or are they worse educated? That is what we need to get to. That is where I think comparability does not really answer the question because it is not going to tell us that, partly because in order to find out whether people are better educated for today that is going to mean slightly different criteria.

CHAIRMAN: I am going to slide over what constitutes “better”.

CHRISTINE FRACZEK: My name is Christine Fraczek. I am a project manager with the Scottish Qualifications Authority, if you have not already guessed! Just to say that in terms of making data more accessible to the public, to our candidates, to our practitioners, I think it is something that we are engaged in very, very strongly. If you look at any of the websites of any of the awarding bodies you will see more and more guidance. Guidance that we give to the markers, to the examiners, the principal assessors, how we set grade boundaries. It is all out there. Where I would say that it is a work-in-progress is that we have to give - without being patronizing because it is quite a technical area - more help in explaining the relevance of the data. For example, the Scottish Qualifications Authority provides the Scottish Government with masses of attainment data which is broken down by all sorts of characteristics, background, ethnicity. This is a wonderful publication that comes out every year but the data goes from the Scottish Qualifications Authority into the Scottish Government and it is making the loops back to the Authority, to the awarding bodies so that we have produced the data and we then use it to effect change. I think it is a work-in-progress. I think more guidance is required and also I would accept that perhaps where we are slightly weak, and again I have to focus on this, is in analysing the data. You can produce uptake statistics and all the awarding bodies do. If you look, we all have our Excel spreadsheets and you will see masses and masses of attainment and performance data, but it is not everybody’s cup of tea, to be honest, and just putting that into more plain English, “Here is the data; this is what we think it means; come back and comment, and how can we make it clearer?” That is the point.

CHAIRMAN: Thank you. There was a lady in the middle there.

TANDI CLAUSEN-MAY: Tandi Clausen-May and I do not belong to anybody, I am independent. The public understanding of assessment, I think Roger Murphy said that it is not that good and if it were, then we would actually have a problem. If the public understood that if the same student took the same test on a different day, let alone a different test in a different year, they would be likely to get a different result, if the public really understood that, then we would lose their faith and I think an awful lot of things would come crashing down. People would get very cross because they did not get into university and if they had happened to take the test on a different day, they probably would have got into university. I do not think we can afford to allow the public to really understand what we do.

CHAIRMAN: A fair point. Simone Aspis, your comment is relevant here.

SIMONE ASPIS: My name is Simone Aspis. I am a PhD student in education at King's College and I am also the Alliance for Inclusive Education's campaigns and policy co-ordinator. I am interested in the whole debate around standards today. I wonder whether the issue is not about the actual standards themselves, but more to do with the artificial values that we put on to these standards, hence the brands of qualifications, NVQs, BETEC, GCSEs, et cetera, and the grades we give to particular standards. Is there anything we can learn from something like the Duke of Edinburgh Award where everybody attains differently but everybody leaves with the same award, everybody achieves at a different standard? I just wondered whether there is anything you think could be learned from that as well.

CHAIRMAN: The gentleman in the middle followed by the lady at the back.

HEW KYFFIN: Hello. I am Hew Kyffin from 5S Consulting, mainly involved in occupational standards but I am very interested in today's debate because I am currently researching maths and what happens to people with maths qualifications after they leave school. I am picking up from the question that you asked and I handed one in, which was Professor Murphy's comment that two people getting a C at GCSE maths could have had very, very different experiences. I do accept that we are talking about approximations here, but that sounds like a lot too approximate, to be honest. We know of a very prominent official within the Sector Skills Councils who recently said employers have "no confidence" in the examination system. I think he was talking about GCSEs, A levels and degrees. I hope that was a bit of an exaggeration for the sake of effect because I think the situation is much more nuanced and complicated. If we have a system where people come out, as it were, with some sort of stamp on their foreheads, there ought to be at least a certain amount of commonality of achievement reflected by whatever has been stamped on their foreheads. I would say that something like the variability between, for example, a C grade at GCSE maths, a particularly good example, is far too great, and that is one of the things I would have thought the system does need to look at.

HILLARY PHELPS: Hillary Phelps, Wycliffe College. I am a teacher. I question the notion that the same student would get a different grade in many cases if they did the exam on another day. I teach psychology A level and I have taught it for near on ten years now and I think the exams are quite cleverly designed to make sure they cover the whole syllabus. Certainly in the upper ability range the youngsters I have taught will master the whole syllabus. Whatever questions they get they will be able to answer satisfactorily.

MARION GIBBS: Marion Gibbs. I am a head teacher at James Allen's Girls' School in London and I did put a question in which comes up here. I am so struck by the stress that the whole system is causing in schools to teachers and students and the fact that we have now got an A\* is another hoop to jump through. It is ridiculous to say that curriculum sets exams. I have been teaching about 35 years (I am just wearing well!) and I have been a chief examiner, all sorts of other things in my life, we all know that what happens is the new syllabuses come out, specimen papers come out and everything changes. I would love to mention the word 'commercial' but this is not the moment in the debate to say that. I think it would be wonderful if someone were brave enough to say we cannot make exams have standards which are multi-purpose. Higher education wants something very different to employers and pupils, learners, students, teachers, parents want to enjoy doing what they are doing in school, having

that validated and not just ending up about jumping through a few hoops on a tick sheet, as it were, and the rest of anything they have done not being validated, which then means that in some schools you would end up reducing, as we have said.

Can I tell one little anecdote. We send lots of sixth formers into local state primary schools as part of their extended curriculum in Lambeth and Southwark. They do enrichment, they do music, drama, sport, all sorts of lovely things, and we have just received the first letters from the head teachers saying, "Next year could your girls come in and coach them for the Key Stage 2 SATS; we feel this would be a better use."

CHAIRMAN: Just before we move to Mike, Sarah from Warwickshire says: "I love learning and I am a complete academic, however at school there is no time for me to enjoy the subjects as there is a ridiculously heavy and extensively prescribed syllabus that I am forced to cram in order to sit an exam. My exams are in six weeks and I still have topics to be taught. There should be more emphasis on debating, discussions and lessons rather than the race that it feels there is to learn everything for the next exam, a cycle which is stressful and I am sure, as you have pointed out, uninspiring".

MIKE BAKER: Mike Baker, columnist for the *Guardian* and the BBC. Two questions really: firstly, we have not heard anything about comparing standards internationally so far today. If it is so difficult to compare standards over time, is it more valid or is it more valuable to compare standards across different countries? The other question relates to the fact that all of the speakers seem to have in one way or another put up a sign which says 'politicians keep out', but if the politicians keep out of qualification reform, who is to decide? Is it the teaching profession, in which case is it the NUT or the NASUWT or the ATL, they may not all agree, is it academics, is it employers, it is the exam boards? There are so many people who could have a say and it could end up in an even bigger muddle than having the politicians in charge.

CHAIRMAN: We got two points there. Paul Muir, you wanted to come in on much the same question. If we spend a little bit of time on international comparators, I think, and then possibly come back to who ought to be running it.

PAUL MUIR: Paul Muir from PA Consulting. We have been doing some work in Saudi Arabia and Dubai looking at curriculum reform and one of the things they use to look at comparators and standards is things like PISA and TIMSS. My point is are they not good comparators for us to look at in terms of looking at standards across countries and places like Finland and South Korea, which are always held up as best practice and doing well? Is that something we can look at in terms of standards?

CHAIRMAN: We have a lot of people who understand the technical side of things in the audience. Does any one of them want to pick that one up, or would you all prefer to move to Geoff Lucas who is here from HMC, who is suggesting norm referencing as a way out of all of these problems?

GEOFF LUCAS: I simply wondered what the panel's view is or, indeed, anyone who is here, given the technical difficulties we have had talking about maintaining standards over time between different subjects, between different qualifications, given that there is very much a lay perspective to this, how can the public understand a very complex system, would there be some advantages in moving back to a system where the top ten per cent in any year, in any given subject get Grade A, the next 10 per cent get Grade B, et cetera. More focus, attention,

time and effort would be put into getting the specifications right, getting the question papers right and crucially getting the marking right. It would take all the effort away from the awarding process, which would be pretty automatic. It would be publicly intelligible. I know the arguments against, such as that one year, the 1986 vintage, might be totally different from the 2010, but I really do not think that is any more difficult to understand - in fact, it is easier - than the whole business of comparability over time and the other issues we have talked about. Is it time to re-open that?

CHAIRMAN: As one who has tried to explain the UMS system, I would sort of agree perhaps.

BEN JONES: Ben Jones, Assessments and Qualifications Alliance, AQA. I would like to pick up on that last point. The first thing I would like to say is that we talk about the A level standard, as far as it exists the standard is designed at subject level and yet the decisions made to feed into that standard are made at unit level. Nobody knows what an A level standard looks like because the examiners, the chairs just look at the unit. Moreover, there is plenty of research now showing that even experienced experts find it very difficult to identify a standard at unit level compared with the previous year. In fact, nowadays with on-line awarding they do not even mark whole questions. It is going to be increasingly difficult for this expert judgment, which the whole system seems to be based on, to maintain the standards. Certainly AQA's view is, especially with much more individual candidate level prior attainment data and so forth available, that the award should be more statistically driven, not necessarily norm referenced like you described but maybe cohort referenced, so you take into account changes in the ability of a cohort. It would stop grade inflation and provide a more transparent way of setting standards rather than having this huge industry of judgment which people think is working and when you analysis it, it may be a little bit flaky.

PAUL NEWTON: Paul Newton, Cambridge Assessment. The cohort referencing is very attractive but the issue is not really standards over time, it is standards between boards. If you do norm referencing between boards and one board has got a much better intake, then it is obviously fair and you really cannot go down that route. I think that is the main challenge. It is worth also saying that there has never, ever been a time in the history of UK examining when norm referencing has been done in the way that most people think it has been done. It is very easy to establish that by looking back at the pass rates back in 1952 when the A level was introduced but you can go back further. You can go back to the school certificate and the higher school certificate and look at the pass rates there. We have often done something that looks like norm referencing, keeping a fixed percentage of the cohort awarded the same grade, but it is done for a slightly different reason. You are just assuming that the cohort is similar, so you award similar pass rates. It is not actually keeping the pass rates stable for the sake of it and that is the important conceptual distinction you need to make there, but I have sympathy. It would all be easier if we could cohort reference and norm reference.

CHAIRMAN: It seems to me the debate is already splitting into people who are interested in learning and education and people who are interested in the arcania of examining, and I am not sure if we are going to be able to bring those two together and, indeed, perhaps part of the problem is that those two are not as together as they should be. Since we are about halfway in we might take a quick round-up from the panel on any thoughts they have had so far because we have covered three or four topics now. Roger, would you like to start?

PROFESSOR ROGER MURPHY: I think one of the problems in our debate is there are a lot of people in this room who are much more interested in improving learning and improving achievement and less interested in assessment unless it can help for that to happen. I think the kind of route we go down with this last point, if you go to norm referencing and you get around Paul's problem by only having one exam board, you can publish the results before the students take the exam, you know what they are going to be, it is fixed in advance, so it will tell you nothing about standards over time. Fine, that would be one good outcome because then you would have to go looking elsewhere. I am interested in Mike Baker's point about international comparisons. Presumably we are interested in looking internationally to find out if people in other countries have found better ways of teaching than we have, but I do not think spending lots of time staring at exam results is the thing to be doing. It is much more interesting to look at videos of how people are teaching in countries where we know the students are achieving quite highly. I think we have got a polarised debate and most of us are more interested in learning and good education. I think the answer is that exams are never going to help us with that, they are not going to answer the questions we want to ask about how do we make learning more effective. We have got to look at other kinds of data and other kinds of evidence.

PROFESSOR GORDON STOBART: I pick up on Mike Baker's politics side of it that if a government is funding 80-85 per cent of what goes on in schools, they have some rights to ask, "What is going on here?" The issue becomes to what extent you control it, whether you hold the ring or whether you get in there and start to manipulate things that are going on. I think that is the current political dilemma at the moment. I think it links to the whole issue of international standards as well because what has happened is these are low-stakes tests for the kids who take them but they have become very high stakes for government. You have governments around the world changing curricula. Denmark and Norway, who have performed badly in PISA - "we have had a bad PISA" - have started changing things radically. Then we go off and we do policy borrowing. We see what has happened in a country that has done well. I found Finland particularly interesting. Having won PISA, our inspectors, and everybody else, went off in the usual coach load to find out what happened there and they found the reason might have been that they had local comprehensive schools with highly paid teachers, who were much respected by the community, and they were keeping libraries open, so everybody just shut up and came home really because we have just got rid of all that! It is the way these link together. Government feel that for them where they are in these tables is important and the political issue is where do I interfere or where do I try and control and where do I back off and leave it, and do we trust the professionals or do we not? In the present election this is quite a debate at the moment, it seems to me, or a subtext of what is going on here: do we trust the teachers and the schools and the exam boards or do we not?

ANASTASIA DE WAAL: I think international comparisons really are the only thing that we have today that is useful. There are lots of caveats. Sometimes we might do very badly for a completely different reason or quite often we might do very well for a completely different reason. The point about international comparisons is they are not testing what we have learnt to learn in order to be tested in. That is the problem with exams at the moment. It is too much about how good you are at preparing kids for those exams. The question is how do you overcome that? Probably the best way to try and overcome that is to make sure that there is a focus on the curriculum and not the exams. That is why I do think to an extent this focus on exams is something of a red herring. It is very important to know how we are doing for kids themselves, it is important for the next thing they are going on to, and it is crucially important

for teachers, but it is this problem about trying to sever national accountability from individual student performance. That is the thing I think we are not quite getting to the heart of. Generally I think that ensuring impartiality, what Mike was saying about politics being involved, we do need to have something which is like Ofsted but better (and Ofqual should be the answer) which is about regulating the details of exams and of curricula as well. I think the big problem at the moment is that we are far too interested in the headline results. We are not interested in details and I think that is why the academy issue is important because a lot of academies that were doing fantastically and were transparent about what they were doing were very proud to show us about that. It is the lack of transparency because we are only interested in the currency, the grades, which is where you get into problems.

CHAIRMAN: Tim, I have left you till last partly because you have spoken quite a lot about the alignment of education and assessment, but you may not want to speak about that.

TIM OATES: I will pick that up first. I want to really address two points and integrate a number of the points in so doing. The second point I want to move on to is very much this issue that Tina raised in terms of the accountability of the system to the public and the social justice inherent in assessment. Absolutely there is this issue of the primacy of the educational experience and the role of measurement within that. We are extremely concerned as researchers, developers and awarding bodies when it is evidently the case, when one looks at the research and goes into schools, that the performativity agenda is dominating, and that the role of qualifications and the role of assessment has been transformed from fair and accurate assessment of what people can do and have done to an agenda for narrow drilling towards assessments. That transformation is not a good thing. Within that there are smaller aspects of the implications of those changes on which one could focus, like the borderline C/D issue, where getting a C is so crucial that there has been an inappropriate degree of targeting on the borderline D/C candidates, to the neglect of both the least able and the most able. Those are narrow technical preoccupations but they actually have profound educational implications in the schools. So the role of assessment undoubtedly should shift to being that where we provide fair and accurate assessment of rich learning.

To pick up this crucial point about public understanding and accountability in respect of judgment, and coming on to Paul's point about are we in a world of criterion referencing or norm referencing? Surely, it cannot be the case that we operate in secrecy and without any degree of moral accountability. We are awash with data in awarding bodies and we use that data in order to inform the judgment process. A number of people in the assessment community believe fervently that what we are about is bringing the right people together to judge what a young person has done, and that is what should be the basis of the fair and accurate assessment of an individual or a group of individuals: the right judgment by the right people. Of course, there are approximations in that. That judgment will be volatile over time, but the key thing is that the right group of people are involved in making those judgments, informed by evidence.

Where it becomes problematic is how much evidence and where should that evidence be derived from? Notionally the evidence should be scripts. It should be evidence created by the young people themselves. What we are experiencing at the moment is quite a subtle and unrecognised shift towards statistics, where people's attainment is being judged not by what they can do but what has been done before. We are neither in a world of norm referencing nor in a world of criterion referencing, and there are big social justice issues involved in that. The drive towards statistics cannot account for an unexpected and dramatic acceleration in

achievement in a given group by virtue of a school or a group of schools doing particularly well. That is a debate which we feel should be worked out in public.

In terms of data we have the data. We need to decide publicly what data we want to get out there through the regulator so that awarding bodies can be held to account by the public and by key institutions.

Judgment is at the heart of it and I agree entirely with Roger. The main thing about awarding is that we need to make sure that the right people are there at the centre from the subject community making those judgments. Just a footnote, Chair, just bear with me for just a second. There is evidence in terms of whether standards have slipped in particular subjects, in terms of how well young people are being prepared for higher education, or how well they are being prepared for employment. There is evidence out there. What is interesting is that that evidence tends to be in specific subjects and it points to the need to make very specific changes in respect of the curriculum and assessment in specific subjects. It does not point to constant, wholesale system change or moving the proportion of assessment, the amount of time spent on assessment, in all subjects from one quantum to another. What it points to is careful and sophisticated management, at a subject level, of the battery of qualifications that we have, and drawing evidence from research, from subject communities, professional communities, and from the economy to manage those changes.

CHAIRMAN: Gordon, you wanted a very quick riposte?

PROFESSOR GORDON STOBART: Just a quick comment. I think we can tweak exams, we can make them fairer and everything else, and I think we have a system that is very conscientious about trying to do this. It seems to me from the other comments it is actually the accountability agenda that should be the prime target of reform. We should be discussing whether we can make schools accountable in other ways than the raw grades? That is what is driving it. It is not the results; it is the uses the results are put to. That is where the political debate ought to go about whether we can back off the accountability and find a better way of doing this.

CHAIRMAN: So you are beginning to answer your question on where the state comes in. Dr Ron McLone?

RON McLONE: Ron McLone, Cambridge Assessment, retired, otherwise of no fixed address! A lot has been said of the real underlying issues that are around our debate and particularly, I think, what is coming out is the underlying purpose of what it is all about. It has seemed to me over many, many years we have seen governments, awarding bodies, employers (we have not mentioned employers much this morning) all having views about what these examinations are about. How about thinking about what the students need? It seems to me our problem is the multifarious way we want the same assessments to actually deliver. We have mentioned that several times over this morning but do not actually get round trying to solve it. We are faced with selecting for universities. We are faced with employers who have requirements which are really brought together sometimes by groups of employers, but actually employers saying, "For my purpose, this is not right," and you also have, of course, schools and teachers who have their own needs and their own wants for the children they have and they know what is best for their children, as schools do. And then we have government. On top of that we have what are called the measurement for accountability purposes. We are awash with data. We do not need any more data, I would suggest.

However, we do need to know how to interpret data, and that is the biggest problem we have because individuals have different ways of interpreting it.

What I would say about the whole way in which we are moving is that because government feels it wants to use what we have got in terms of the results, we are enforced to use the same instruments to measure too many different things. We should concentrate on separating out what is needed for accountability from what is needed for the individual students so that the student is best prepared for what they want to do in the future and at the same time get away from that measurement which implies there are equivalences. Part of the problem we have also is that because we think we need to measure, we think things are equivalent. You know the sort of thing. We have mentioned equivalences between subjects that are somewhat false. We talk about equivalence between qualifications so that you can actually get schools, for example, or academies, using IT qualifications that count as four GCSEs and regarding them as the same as four different subjects like history or geography. We have to ask ourselves why do we do that? Is it because it is most important that the state needs to know, or should we really be valuing that it is most important that you develop individuals to the best of their abilities so that they are best certified for what they can do. I happen to think it is the latter.

**HELEN PATRICK:** Helen Patrick, retired. I owe this comment to John Bell, who is nothing if not a statistician. I want to go back to the point about norm referencing and other ways that might get rid of the complexity of grading and judgment. The risk of going down the statistical route is that you divorce the result that someone gets from the content of what they have learned and the performance they have exhibited in terms of whatever kind of assessment they have done. I think that going down solely the statistical route is not a good idea.

**CHAIRMAN:** We have not said very much about employers but Ron is anticipating me. We have somebody from the Manufacturing Technologies Association at the very back and also I am going to call David Delmont to talk about: is this just GCSEs and A levels? Does it impact on vocationals?

**PAUL O'DONNELL:** Paul O'Donnell from the Manufacturing Technologies Association. I was heartened by Ron McLone's comment about the need to involve employers more. Somebody else earlier said that employers have no confidence in the examination system. I think that is probably over-stating the case. The issue is actually not whether they have any confidence in it but what their engagement level is in it. I think there are two underlying problems with employer engagement. The first is that events like this are very welcome, and I thank Cambridge Assessment for putting this sort of thing on, but they do not necessarily speak to a wide range of audiences, the people who are effectively the customers of the system, and that means employers, that means parents and that probably means students as well. The other underlying issue is around the involvement of employers in the development of qualifications. I wondered whether the panel or other people in the audience had anything to say about the mechanisms that are available for employers to be involved in development and whether what we have got now is working, which it does not seem to be, and perhaps how we might want to improve that in the future.

**DAVID DELMONT:** David Delmont. I am a worked-based learning manager in a further education college in what is the fairly deprived borough, Barking and Dagenham. I have been there for six years and it is my view that the standards of the young people applying to us for apprenticeships have dropped. What the young people tell us is that when they are in school,

if they are going to be taking the vocational route rather than the academic route, ie going from GCSEs on to A level, then on to university, they are saying they feel they are, I would not say sidelined, but they are on a different track and they are being ignored and they are not getting the same standard of education as the people who are taking the A levels.

Every year in August we look and see the young people opening up their GCSE results and their A level results and that is what seems to be the measure, but we have got people who are taking a vocational route and, quite honestly, they are coming to us with very rarely grades A to C and often no GCSE grades at all and often very low grades. When we take them through their qualification, the further education college ends up as a surrogate for what they have not learned in the previous 11 years.

An apprenticeship is not an easy option and I think sometimes there is too much emphasis on what is seen as the practical side. An apprenticeship contains three main parts: it is a technical certificate, an NVQ and also key skills, which now can be developed into functional skills. We find it extremely difficult to achieve apprenticeship frameworks on the key skills side because the young people have not got those skills coming to us. Although we give them additional learning support, it is extremely difficult to get them through. The Government says it wants 500,000 apprenticeships. I am hoping it wants 500,000 apprenticeships to pass their frameworks. It is not just going to be a statistic that we have got enough people coming in on apprenticeships; it is also important that they pass. Further education colleges are measured by minimum levels of performance. If they do not get learners through their qualification, then they will say next year, "You have not achieved your minimum level of performance, therefore we will take the funding away." What we are really looking for is when young people apply for apprenticeships, they are equipped with the functional skills to achieve their apprenticeships because it seems to me that we may be turning out young people from schools who are functionally, I am not saying they are illiterate but they may be bordering on that. Thank you.

CHAIRMAN: Thank you. The gentleman in front of him.

DAVID HOWE: Hello, David Howe again, but can I wear a slightly different hat as a governor of a new institution starting in September called the School of Communication Arts. I will come on to why that is relevant in a minute in terms of its employer links. I also wanted to comment on two issues which have come up in the last five minutes and one is this whole question about are we asking the right kind of questions and the right values in this debate and this thing about academies or other schools that might be doing BTECs in order to raise their qualifications. I am not suggesting this is necessarily the case, but as a senior manager possibly you might be playing a very clever game by doing that, of saying if you believe that all this debate about whether standards have fallen and examinations is not about the right kind of questions, then, okay, let us do these BTEC examinations, drive up the results because we have established an academy where there may have been (so said) a failing school there before, to get rid of that debate so that the politicians will go away from us so we can concentrate on what the more important questions are, which is about what we should be delivering and what we should be teaching in there. If you are a senior manager you might take that view, what is the quickest way to get the political pressure off us and stop the kind of criticism and that could be one view then.

Can I pass comment about another area which has been raised just recently which is about employer engagement. It comes back to the School of Communication Arts, which is around

the arts and communication industries. It is a new specialised institution which will deliver very vocational qualifications. It is obviously quite small and it will be starting in September. What they have done there in establishing the curriculum is they have put it out - and I am sure there will be professionals in the audience who will be horrified in listening to this - on Wikipedia asking about how to do it. They passed it back to the industry experts and said, "Okay, you quite often say we are not preparing young people for the right kinds of things, so let us pass the buck back to you and say would you like to contribute directly in to the curriculum?". They invited people from the industry but also it was open to other people as well to design the curriculum on-line and to contribute to it, to invent new models, to change the content and to revise what people had put there. Obviously it was moderated in terms of getting it up to some consistency, but it was about putting that back out to the industry and saying, "Okay, the majority of the input into that curriculum is going to come from the industry, not from the educational professionals, and let us see where that comes from."

As a method of doing that right across the wider state education system, that is obviously an impossible thing to do and to manage, but it does raise a rather interesting way of doing that in a specialised area that is probably worth looking at in some detail a couple of years down the line to see what that achieves and whether the people who come out from that school are seen by the industry as being more relevant and more trained in terms of the kinds of skills they have got.

CHAIRMAN: Thank you for that.

MARTIN ROWLAND: My name is Martin Rowland and I am freelance. I have worked in further education, higher education, I do some work with sector skills councils and I work with an examining board as well, so quite a range of experience. I would like to reiterate what the previous speaker was saying. In the work I do with the Sector Skills Council it strikes me that there is a dichotomy between the employer side and the education side, and I think it reflects back to what Roger Murphy was saying earlier on and several others, unless we decide what the fitness of purpose is there is no real point in talking about standards because industry employers want one thing and educationalists want another and where is the standard in that? It leads in to the role of government, which I think is absolutely key because even if it does not regulate, I think it is probably the only medium that can get the interested stakeholders together to decide what the purpose is.

A second issue which worries me a little bit from my work with exam boards and working with schools and colleges in a different capacity is that ultimately I do not think a lot of our learners give a monkey's about standards and I will tell you why. A level students are using the data which is out there to inform the decision about what A levels to take and what I would like, what interests me and where it will take me in terms of a career are getting more and more irrelevant in the choices that young people are making. They are simply looking at the raw statistics and saying, "You will get more Grade As in this subject than you do in that. I stand a better chance of getting four Grade As if I do this, this, this and this than anything else." I think that is probably one of the most worrying things about the standards debate. People are going nowhere because they do not know their purpose, the exam system is going nowhere because it does not know its purpose, and until we have got that worked out, there are no standards at all.

CHAIRMAN: Thank you. I seem to remember reading some kind of research paper on exactly this topic. Sylvia Green on the first part?

SYLVIA GREEN: Sylvia Green, Director of Research at Cambridge Assessment. I leap up at this point because we have conducted some very detailed research into A level choice, which some of you may have seen, and I would like to portray a slightly more optimistic picture of our A level students because it was a big study, it included about 12,500 students, I think, and it is available on the Cambridge Assessment website if anyone would like to look at it in detail. What we found with the students interviewed was that they were making choices for very good reasons. They were making choices about subjects that they enjoyed and subjects they were interested in. Yes, that very often is the subject in which they may have done well, but they were not looking at it in that particular way, I am pleased to say, from this research. They were looking quite clearly at the kind of futures they wanted for themselves and the kind of professions they might want to pursue in the future, but they were making good choices for - I felt and we felt - very good and useful reasons. I would like to point people towards that research, if anyone is interested in A level choice and they were getting advice from lots of different sources, perhaps that advice could have been better timed and better placed, but the students themselves were pretty savvy about what they were doing in terms of their own subjects and their own futures. We were very, very pleased with the findings from that research; it was an optimistic finding.

CHAIRMAN: Thank you. That deals with the students but it does not go to the heart of your question as to whether the system knows what it wants to do. It is interesting that the 1911 report that I referred to earlier recommended a certificate in which “the standard shall be set at the level of a good liberal arts education at age 16”. Do we have too many targets, too many standards? That seems a very simple standard. Whether it is understandable in modern Britain is another matter, but it is a simple one-off standard - pass, fail, distinction. We could go down that route. Who would like to pick that up?

JULIAN STANLEY: Julian Stanley, University of Warwick. I am sorry, it was not exactly that point I wanted to come in on but I was trying to come in on an earlier point for some time. I wanted to bring the thinking back to the question of public understanding, which we talked about a lot, and what is wrong with public understanding. Again and again people have been talking about this gap and there does seem to be this persistent public understanding of qualifications as differentiating. That does seem to be one of the problems why public understanding persistently sees something normative going on there and persistently is dissatisfied if that expectation is not met, if too many people do well basically. That may be something we can educate on, but that may be a basic public understanding which is out there which has got to be come to terms with. Public understanding does seem to be not entirely uniform. There is a notion out there that it should be fair and, again, that is something which technically we find quite difficult to deliver. Again, there is a public understanding which is that qualifications should tell us what people are going to be able to do in the future, a sort of validity. I suppose it is getting the public to think about those different claims and try to reconcile these different expectations they have got of qualifications and getting different bits of the public to explore that a bit more might help the public match up their expectations with what they are actually getting.

CHAIRMAN: Thank you. I see Tim wants to come in. Was it on this point?

TIM OATES: I want to return to the vocational point, if that is okay.

CHAIRMAN: Let us stay with vocational for a bit.

TIM OATES: In terms of international work, Mike, German researchers and policymakers just look at you as if you are crazy if you say, “Of course you have got the same number of modules in all of your occupations, haven’t you?” or, “All occupations need to achieve the same level of functional skills.” They just think you are mad. They have a tradition of adapting the requirements of the qualification system to the requirements of each industrial sector. The timeframe for refinement and modification of the content is entirely in tune with the clock which is ticking in each sector and that is very important.

There are examples of very, very good practice emerging in our industrial training sector, and Linda Clark and Chris Winch have done a great deal of work on construction where things are working very well. What is interesting, though, and it gives us some strong signals, is there is a degree of self-organisation amongst employers, there is a very strong link between the qualifications and work processes and there is a link between qualification and licence to practise. That all makes tremendous sense. What is difficult about construction is that when you listen to employers - employers not employer organisations - what they say is, yes, we have managed to make sense of the latest qualifications which have been given to us in terms of our work processes and it is now working nicely for us. They have taken diplomas and combined it with NVQs and worked with FE colleges to get sensible programmes in place, but is it not a problem that it is that way around, that what they have are comprehensive learning programmes tuned to work processes and they have to adapt the national qualifications to those? In other words, we should have a responsive qualification system, not one which enjoys that sort of relationship with work.

RON McLONE: Could I support Tim in that. It seems to me that what we tend to do is because of adapting the notion of fairness (and I understand the fairness) that we equate that with equality and then translate equality into equivalence. Equivalence then means that we believe that we have to have a common system that the end users have to adapt to. That is quite a different approach to one which might start by saying what we really want is a system which is responsive to where the end users want you, and if that means that your standard-setting through the curriculum and through attainment has to be differential, then we have to live with it. The problem is that is the debate we have not been having.

CHAIRMAN: That seems to have got a lot of heads nodding. Anastasia?

ANASTASIA DE WAAL: Just on the employer relationship and the vocational qualifications, crucially, I think that there is a very bogus commitment both on the part of the Government and, sadly, on the part of many examining boards about vocational qualifications. I think it is being utilised as a good way of being able to ratchet up achievement by using what are not vocational qualifications. I am particularly talking about GCSE equivalents. We are talking about pseudo vocational; they are not really vocational, they are pseudo-academic. We know very well from a lot of employers, to take construction as an example, what they are saying is, “We do prefer Eastern Europeans actually because when they come into construction they have a solid foundation in geography. They know if we send them to Birmingham where Birmingham is, for example. What we do not want is somebody who has done a GCSE equivalent in construction who has learnt a bit about construction or what people in construction do, which is really not very useful.” I think that is highly problematic, not least because it is depriving a lot of kids of a solid foundation. Why? Because we think it is better for them to achieve a higher grade in something that is not terribly useful than a lower grade in something that is useful but, crucially, because I think it

is really undermining vocational work. It is saying that weaker pupils go into vocational qualifications. It is saying this is what vocational work is about when it is absolutely not vocational. It is also saying that in order to succeed in vocational work you do not need any solid academic skills. That is a big mistake and if we look at international examples where they are successful in terms of both status and in terms of quality, that really is not the line of reasoning there.

SIMONE ASPIS: Again I will ask the question after the gentleman who just iterated the point I was making. The Duke of Edinburgh Award is a very highly regarded award. Lots of people rate it: employers, educationalists, schools, students alike. Everybody attains at a different level but everybody leaves with the same award. So for example the skills section of a Duke of Edinburgh Award, you can develop an academic skill or you can develop a vocational skill. Every student learns at a different level. For example, you could have two students who may want to develop their French skills. One may be able to translate up to the highest level. One may only learn a few words in French. Everybody achieves differently but everybody leaves with the same award as long as you participate in the other sections as well. The important thing is there are no prescribed standards. What can we learn from this? I kind of feel that we have not actually even acknowledged that there are some lessons that we can learn from that because, remember, the Duke of Edinburgh Award is highly regarded. As far as I can see, nobody has said the standards have gone down, unlike the other qualifications, and I suspect that is because we are too busy branding, rebranding, putting different grades to changing content, unlike the Duke of Edinburgh Award where actually it is the individual who drives their programme. Each individual designs their programme and each will attain differently but equally, so it is really important. I really would love to have a response on that. Thank you.

CHAIRMAN: I am going to tell the Panel I would like them to address that when they do their winding up remarks. It raises interesting questions, one of which is how long would it last if it was actually a passport to university.

TESSA STONE: Tessa Stone, Chief Executive of BrightsideUNIAID, an educational charity. I wanted to come back to the point that Tim and various others have made about the proper relationship between business, higher education and the exam creation systems, because one of the unintended consequences of getting that relationship in the wrong order is that you see HE and businesses going off and doing their own thing. We have talked quite a lot today and we have heard quite a lot today about the professionalism at the heart of our exam standards, which I think is fair. One of the things that concerns me is that when the exam system is not seen to be fit for purpose, either for higher education or for business access, you then get a sprouting of unprofessionalism, or people trying to adjust that who are not experts in this area.

Two specific examples. As an ex-HE admissions tutor I know that you introduce additional examination for HE access at your peril. There is a considerable social justice question to be raised about the growth of additional exams for entry to HE. They are vastly differential and certainly when I was involved in creating additional tests we were not doing it from the point of view of experts in testing. We may have been experts in our subjects but we did not know about fair testing for that sort of outcome.

Likewise, in the interesting example at the moment of the access to the professions debate, where because there is seen to be a lack of a fair route into various professions, the professions are now being brought in at the end to say how they can adjust this, "We do not

understand or we cannot use A levels, therefore we will set up an additional set of criteria from the basis of our experience,” and as part of all the discussions I have had, with the best will in the world and from the best of intentions, these people are not necessarily best placed to implement some of the things that they are intending and they are not best placed to put this in the context of a wider education debate, and that really does concern me.

CHAIRMAN: Thank you. I am going to take one more question from the back and then I am going to ask the panel to give us about two or three minutes just summing up, just commenting on what has happened in the debate or if you want to make your point again, by all means do so.

CHRISTINE FRACZEK: Two very quick points. I think the last two speakers are talking, as you all know, in standards speak about the qualification standards. That is where we define the knowledge and skills and, with respect, I think that is the starting point because that assessment gives an opportunity to achieve the qualification standards, so I look at that as the end of the process. The heart of it is how do we define the knowledge and skills in the qualifications themselves and then how do we go about proving, if you like, that learners have achieved them - I think that is another very interesting debate.

The second point is something that Tim Oates mentioned earlier, that at the end of the day statistics are all very well but it is the candidate scripts. Speaking as just one awarding body, I know it happens in Scotland, and it certainly happens in the rest of the United Kingdom, and you will know about it, there has been a programme of archiving candidate evidence over a long period. I am sure that Ofqual has taken this on from QCA, it is a legacy programme, and that allows the awarding body to do (and I will be honest) on a limited sampling basis a comparison over time.

The last point I would like to say is certainly in SQA, and I am sure the awarding bodies in England would share this, we are all greatly encouraged by the advent of digital evidence. You now have e-portfolios and e-scripts, not to the same extent, but certainly I am involved in this programme, and I am looking at instead of having to get huge big folders in to keep from Scottish vocational qualifications, we have run a pilot where we will now have e-evidence. Just for the logistics of it, because there is a cost to store lots and lots of candidate evidence, that programme has been ongoing and, as I say, personally I am encouraged and hopeful that we will get much more efficient by having more e-evidence. Having spoken, though, to the Archive of Scotland, I think there is an issue about how long digital evidence will last. I know we have had paper for centuries, but that is yet another debate perhaps.

CHAIRMAN: Thank you. I am going to turn to our Panel and ask them to give us a final round-up of thoughts. I am going to start at this end with Anastasia De Waal from the think-tank Civitas.

ANASTASIA DE WAAL: For me the clearest point which I feel has been reiterated is that you cannot look at exams in isolation; you have to look at them as part of a process. That is why I do think that a lot of comparing one paper from one year with another year is, in some ways, not actually a very valuable exercise. I think it is very clear that at the moment exams have - and I think it is legitimate to say - been devalued to an extent, and that devaluing has happened through the process. Exams are currently used too often to prove success. In fact, one exam board (I hope it is not Cambridge Assessment!) has a tag-line “Achieving Success”, or something similar, which in itself is very significant, that that is what we are seeing exams

as doing, not being about testing what kids know but actually proving success. The problem ultimately is coming from the centre. From the centre I do mean government, but I also mean other organisations and I do mean at local authority level, and this is having an enormous pressure on schools. What is happening is that the aim is to ratchet up performance, sadly, and not to ratchet up learning, and that is very difficult because obviously the primary purpose initially of ratcheting up performance was to prove that there was more learning going on. This Government in particular had a massive and a very good commitment towards boosting learning and boosting opportunities but perhaps when it became less successful and things were not going quite as rapidly as they had wanted them to, that is when performance rather than actual learning took over. I think in order to sever that relationship again between performance and performativity and get back to performance and learning, we do need to make sure that there is much less opportunity for government to intervene in terms of pressure. Also the market system of examining boards is problematic because, yes, I think it is right to say that exam boards are responding to this need to find weaker qualifications and weaker exams which do allow higher performance, and I think it is fair to say that they are responding rather well, and that is a problem.

PROFESSOR GORDON STOBART: I will just go back to my accountability issue. I think one of the themes that has come out of this is the pressure that exams are under in terms of the use of the results in this accountability system we have here. This must be one of the fiercest accountability systems on the planet. Accountability drives the curriculum and the assessment and the results, so we are in a tighter grip than most countries where they may regulate the curriculum but very few actually regulate all parts of it to quite the same extent. How we can move away from this fierce accountability and all the unintended consequences that we have been hearing about is critical. I think where that leads in things like equivalence is how tightly do we try and make things equivalent or, like the Duke of Edinburgh Award or Tim's point about other countries' vocational qualifications, do we allow ourselves a more flexible system where the currency value of the qualification itself will do the job. We have not talked about things like music grades. Nobody touches those because people know that in a sense there is a standard there and you leave them and you back off. Could we be a bit more relaxed about this in other parts of the system?

PROFESSOR ROGER MURPHY: Two issues. Firstly this public understanding issue. Today for me has been about being a bit more honest in a public setting about some of the shortcomings of exams. I am afraid I cannot go with this view that we should not tell people about this because it is going to be uncomfortable. We live in a modern society where, thankfully, we have a lot more information about a lot of things. I hope I am not being over-dramatic by saying we know much more about poverty and child abuse now than we used to. They are uncomfortable things to hear about but I think it is a good thing that we know about them. Perhaps not in the same category but I would use the same argument - I think we need to know about the shortcomings of exams and assessment and we need to face up to some of the uncomfortable truths about that, and for many of the decisions that we want to make we need to look elsewhere for good evidence.

A second point - the key thing in the debate for me today has been about the relationship between the curriculum and assessment. Assessment should follow curriculum; it should not lead it. A lot of people have talked about good education, improving the standards of education, making education more relevant to young people and adults, allowing people to engage more in the learning process and learn more. That is the important discussion. Assessment then needs to follow that and not destroy it. The history of the relationship

between assessment and education is that far too often it is assessment that messes up education. You get bad backwash effects, you get teachers under too much pressure to teach for a test - that is bad for education.

TIM OATES: As Roger does, I too take heart from the discussion we have had about the relationship between assessment and curriculum, and reinvigorating the relationship between curriculum and assessment is absolutely critical. A couple of points: I will just repeat my emphasis on the need for more openness and honesty about whether elevation in the numbers gaining a particular grade derive from a genuine underlying improvement in attainment in each and every case. Openness and honesty in examination of the evidence relating to that is fundamental. That is why we are having the debate. In talking about change, I am not arguing against change; in many ways I am arguing for a different sort of accountability. What I am arguing is that changing qualifications should be driven by groups such as users and by subject communities, and that makes the qualifications much more accountable, but it is a different sort of accountability. It ensures the qualifications are fit for purpose and are linked into the communities which require them. Of course we should use evidence to modify each and every qualification and each discrete qualification, therefore we need to undertake a systematic inquiry into subjects, into sectors and ensure that we take that evidence into account in terms of the changes which we make. That all points away from the direction of wholesale system change and more towards partnership arrangements in the subjects and sectors in order to drive the change in qualifications. That would bring a new form of accountability to the qualification system.

Finally, a point which we touched on briefly, and it may seem arcane, but in a sense it is opening up a new battleground, and that is the debate between judgments and statistics as the basis for making particular rewards and deciding upon a proportion getting an A or a B in a particular year. That is the fundamental technical debate. In terms of opening up public understanding, that is not a debate we should be having behind closed doors. We need to ensure that the public are involved in that debate, policy communities, user groups and so on, because the accuracy and fairness of the qualifications is a pre-eminent concern of an awarding body and within that we are concerned with the issues of justice which are associated with it. At the moment we have a fundamental tension between groups and organisations, some driving the system more towards statistics and away from judgment and others pushing it in the opposite direction. What we need to add to the debate we have had today is a very open and intelligible debate about the changes which that kind of movement represents.

CHAIRMAN: Thank you. May I point out, this debate will continue for one more month, which will be accessible through the Cambridge Assessment website, so anybody who has not said what they wanted to say or wants to take issue with what has been said, please input into it. We will be taking everything that has happened on the exam standards debate website, together with what we have discussed today, and producing a paper after this, which we will then be sending to the Government, Parliament and stakeholders of one kind or another. I am glad to say that in stark contrast to the normal style of debate, this has been sane, sensible, with much more light than heat. Thank you to our audience on-line. It only remains for me to say thank you to the panel and I will ask you to show your appreciation for them for, in a sense, leading us rather than directing us down the passages we have gone and to say thank you to all of you for coming and to wish you a good afternoon. Thank you. (Applause)