



CAMBRIDGE ASSESSMENT

# Cambridge Assessment Exam standards: the big debate

## Report and recommendations

Autumn 2010

## Participants

### Presenters at the Parliamentary research seminar

Dr. Robert Coe, *CEM Centre at Durham University*  
Dr. Jo-Anne Baird, *University of Bristol*  
Tim Oates, *Group Director of Assessment Research and Development, Cambridge Assessment*

### Panellists at the live debate

Chair – Bene't Steinberg, *Group Director Public Affairs, Cambridge Assessment*  
Tim Oates, *Group Director of Assessment Research and Development, Cambridge Assessment*  
John Bangs, *Assistant Secretary of Education, National Union of Teachers*  
Professor Roger Murphy, *School of Education, Centre for Developing and Evaluating Lifelong Learning, University of Nottingham*  
Anastasia de Waal, *Director of Family and Education at Civitas*  
Professor Gordon Stobart, *Institute of Education, University of London*

### Members of the audience who contributed at the live debate included:

Simone Aspis, *Alliance for Inclusive Education*  
Mike Baker, *education journalist*  
Greg Brooks, *University of Sheffield*  
Tandi Clausen-May  
David Delmont, *Barking and Dagenham FE College*  
Christine Fraczek, *Scottish Qualifications Authority*  
Marion Gibbs, *James Allen's Girls' School*  
Sylvia Green, *Cambridge Assessment*  
Andrew Harland, *Examination Officers' Association*  
David Howe, *School of Communication Arts*  
Tina Isaacs, *Institute of Education*  
Ben Jones, *Assessment and Qualifications Alliance (AQA)*  
Huw Kyffin, *5S Consulting*  
Simon Lebus, *Cambridge Assessment*  
Geoff Lucas, *Headmasters' and Headmistresses' Conference (HMC)*  
Ron McLone, *former Cambridge Assessment*  
Paul Muir, *PA Consulting*  
Paul Newton, *Cambridge Assessment Network*  
Isabel Nisbet, *Ofqual*  
Paul O'Donnell, *Manufacturing Technologies Association*  
Helen Patrick  
Hillary Phelps, *RSA fellow*  
Martin Rowland, *freelance*  
Amanda Spielman, *ARK Schools*  
Julian Stanley, *University of Warwick*  
Tessa Stone, *Brightside UNIAID*

For further details on contributions made by audience members at the live debate, please see the full transcript available on the Cambridge Assessment website at: [www.cambridgeassessment.org.uk](http://www.cambridgeassessment.org.uk)

Comments attributed to individuals in this report are taken from contributions made at the Parliamentary research seminar, the live debate and on the blog, 'Let's talk exam standards'.

---

## Exam standards: the big debate

### Why do we need to debate exam standards?

Over the past few months, Cambridge Assessment has been hosting an open and frank debate to clarify public understanding of the different examination standards issues.

Examination standards – and the perception of them – matter to society and dominate educational and media debates. However, we believed that a comprehensive, unbiased and well-informed public debate on the issue was missing, and indeed long overdue. We also felt that it was important that Cambridge Assessment took a lead on a topic of such central importance to our three examination boards – OCR in the UK, Cambridge International Examinations and Cambridge ESOL.

We urged educationalists everywhere to join a large-scale project that could discuss standards in a sensible and informed manner, away from the fevered and simplistic froth of the UK's annual summer spat. This is a complicated area, and it was important that the debate was well-informed. Therefore, the aim was to get people talking about the same thing at the same time.

### How the campaign was run

There were three strands to the campaign.

The campaign began early in 2010 with a Parliamentary research seminar for leading academics and policymakers at the House of Commons, to start tackling some of the trickiest questions.

England's former Chief Inspector of Schools Baroness Perry chaired the event, which also featured talks from academics who lead in this field.

The second strand was a live debate held on 29 April at the Royal Society for the encouragement of Arts, Manufactures and Commerce attended by over 100 people including teachers, assessment experts, employers and journalists. The debate was streamed live and nearly 1,000 people engaged with the proceedings online.

The campaign was supported through an online debate via a Cambridge Assessment 'Let's talk exam standards' blog. The blog was designed to stimulate discussion before the event as well as being a post-event facility for continuing the debate through to June 2010. To start the online discussion we asked several education experts to comment on a paper written by Tim Oates at Cambridge Assessment about standards in public examinations. The public were then asked to respond to these views on our blog.

The campaign attracted widespread media attention and coverage appeared in the run up to the debate in the Times Educational Supplement, Daily Telegraph, Times Online, The Independent, Daily Mail, Independent on Sunday and Belfast Telegraph. Subsequently coverage appeared after the event in The Guardian, Daily Telegraph, Daily Mail and on BBC News Online.

---

## Report and recommendations

*This paper summarises the outcomes of the discussions, which focused on the following main themes:*

- ❖ What do we mean by standards?
- ❖ Is there 'grade drift' – and what may have contributed to it?
- ❖ How can we increase public understanding?
- ❖ Who should own qualifications?
- ❖ How can employers be engaged?
- ❖ How important should qualifications be and how much can they be relied upon?
- ❖ What are qualifications for?
- ❖ What is the impact of the current accountability system?
- ❖ Do international comparisons help?

## 01 What do we mean by standards?

Given the need for a comprehensive, unbiased and well-informed debate, we began the campaign by clarifying the different types of 'standards':

### Standards of demand

This denotes the assessment challenge. It is what people usually mean when they talk of exams getting 'easier or harder'. It refers to the level of the skills, knowledge and understanding required to successfully complete the assessment.

### Content standards

This is the value or relevance of the content of the examination. So, for example, content standards decline when an examination becomes old-fashioned, redundant or irrelevant to the needs of today. The content of knowledge changes over time in different subjects; the requirements of Higher Education and the economy change; the needs of society change. Qualifications thus need to change over time, in order to maintain their utility and value.

### Standards of attainment

This denotes the results students attain when they take an examination. If successive assessments remain at the same level of demand but the students know more, or are better prepared, then they will attain more. So, for example, their 'standard of attainment' might rise from a grade B to a grade A.

### Standards over time

The issue of 'standards over time' often dominates the public debate. Yet as several of the contributors to our debate pointed out, it has serious limitations and struggles to tell us anything meaningful. As comments on the blog testify, examinations have to change over time as priorities change, technology changes, and knowledge changes. This makes comparing standards over time problematic, as one is no longer comparing like with like.

Roger Murphy argued throughout our debate that there is a problem in comparing current papers with 1952 papers because such crude comparisons do not take into account the significant changes that have occurred in this time. These changes include: more children staying on at school, frequent curriculum changes, smarter exam preparation, and greater flexibility in the way awarding bodies allow candidates to show what they can do. He also argued that our public examination system is not the best way to address standards over time, as this is not its prime purpose. Rather, its prime purpose is to give grades to individuals to help sort out what they are going to do next in terms of their academic or career development.

Referring to research by the think tank Civitas about teachers' views on A levels, Anastasia de Waal said that teachers agreed that comparing standards over time is not a useful exercise. This is particularly true because of the extent of reforms to A levels

and the difficulty of comparing a modular A level to a linear A level. She argued that what people are really interested in is how well educated people are. She said comparing exams in isolation cannot tell us that, as the process as a whole needs to be considered.

Professor Gordon Stobart agreed that the standards over time debate "goes nowhere" and should be abandoned. He compared exam standards with climbing Mount Everest, saying:

*"In 1953 two people got to the top of Everest, an extraordinary achievement at the time. Yet on a single day in 1996, 39 people stood on the summit. That might suggest that Everest had become 20 times easier to climb. Yet the mountain remains the same height. Of course, today people have better equipment, better training, better nutrition and so on. In that sense, it is less surprising that more people can climb Everest. But while that may make the achievement less exceptional, it does not change the 'standard' of the mountain climbing achievement."*

### Standards between subjects

The idea of comparisons between subjects was also raised during the debate. Roger Murphy raised concerns about 'chalk and cheese' discussions, arguing that it is impossible to compare A grades in maths, chemistry and French, as they are such different subjects.

## Recommendation 1

Before any discussion about 'standards', terms need to be defined and clarity reached about what kind of standards are being referred to. In particular, given the widely recognised problems with comparing standards over time, the assessment community questions the value of concentrating on this simplistic argument.



## 02 Is there 'grade drift' – and what may have contributed to it?

The discussions considered whether there had been 'grade drift'. For example, whether a grade A might now be awarded to work that would previously have been awarded a grade B.

There was a broad consensus that grade drift probably existed, although so many confounding factors made it difficult to isolate and identify. How this might have come about was extensively discussed.

### Constant change in the system

One cause was the constant change to qualifications. Tim Oates suggested that *"if you effect continual or inappropriate and unnecessary change of qualifications, it makes holding standards over time extremely difficult"*. He argued that constant change has made it very difficult to maintain standards over time and that in recent years we have seen too many changes in the form and content of qualifications. Whilst qualifications should be updated to reflect changes in each subject, the extent of top-down change has far exceeded the level actually required. More stability is needed to secure high confidence in our public examinations.

One example of 'arbitrary' changes to the examinations system is the way in which politicians have banned, then reintroduced, the use of calculators in maths GCSE. This has changed no fewer than seven times in 15 years. Similar changes are listed in Appendix 1.

Other blog comments pointed out that because changes to assessment structures had occurred at great pace, it is difficult to evaluate what factors have had an impact on supposed changes in standards. A concern was also raised that these changes had eroded public confidence in standards. This, in turn, led to demand for further change, creating a vicious circle.

It was suggested that changes in qualifications should not be top-down but driven by user groups and subject communities as this would make qualifications more accountable. It was also proposed that the UK should move from wholesale system change towards partnership arrangements in order to drive the change in qualifications. Others talked about the 'yearning for continuity' that existed amongst the profession and amongst schools.

### Transparency, accessibility, guidance

Other factors which could have contributed to grade drift included the improvement in the transparency of assessments, which were no longer just *"hitting people hard with things that they do not expect at a time that they do not quite expect it"*. Speakers said the system now ensures that students are given the benefit of the doubt and there is a strong emphasis on the inclusion agenda. Another important factor is that students now have the chance to optimise their progress, for example using feedback from their AS results to make decisions about A2. In addition, because of increased accountability there had been a vast improvement in availability of materials, bringing improvement in terms of efficiency of process, but not necessarily to underlying attainment.

Several participants referred to the impact of modularisation. They argued that teachers were supportive of the idea of a more accessible A level, noting effects such as encouraging boys to work right from the first few weeks of the course, thereby enabling them to attain more and higher grades. Yet, equally, there were concerns about the impact of re-sitting modular exams repeatedly and the effect this might have on a qualification's reputation with universities.

### Recommendation 2

The frequency and scope of change in qualifications needs to be reduced, inappropriate change in the format and content of qualifications should be resisted, and the change agenda should be driven by the user and subject communities.



## 03 How can we increase public understanding?

### Making the data accessible to the public

Several contributors talked about the need for openness and honesty about the limited nature of what can be inferred from an exam result. There was also much discussion around whether the growth in the numbers gaining a grade reflected a genuine underlying improvement in attainment.

The extent to which transparency for the public on this issue is necessary and desirable was also debated. It was pointed out that the pressures for transparency have intensified significantly as qualifications have increasingly been used for accountability measures. However, it was also highlighted that this is a technical subject and there was genuine debate about how far one should go in explaining such detail. For example, is it realistic for the assessment community to aim for full public comprehension of the technical detail of how standards are set and maintained?

Unsurprisingly, the audience struggled with the question of how to use the large amounts of data that currently exist and to explain a highly technical area to the end user in a comprehensible way. There was general consensus that awarding bodies possessed a huge amount of data but that the difficulty was analysing this in a way that the public could understand.

### The judgement vs. statistics debate

Much of the debate focused around the issue of using judgements, rather than a more statistical approach, in discussions about maintaining the standard. Participants referred to the move away from an approach focused on the professional judgement of the quality of candidate performance towards an approach focusing on statistics and data produced prior to the candidate sitting an exam, such as performance in earlier assessment.

Geoff Lucas suggested that given the technical difficulties of maintaining standards over time, there could be an advantage in returning to a system known as 'norm referencing', i.e. where the top ten per cent get an A, the next ten per cent a B, and so on. It was suggested that the benefits of this approach would be to allow the public to understand a very complex system. Some were of the view that given the changes to qualifications, it was going to be increasingly difficult to rely on expert judgement to maintain standards and that instead awards should be based on a statistical approach, where people's attainment is derived from what has been done before.

This was an area also touched on by Dr. Jo-Anne Baird, from the University of Bristol, at the House of Commons seminar. She contrasted statistical and judgemental methods, noting the factors that each takes account of and adjusts for. She presented evidence casting doubt on the ability of experts to consistently make the kind of judgements required of them in some contexts for both setting, and maintaining, standards.

However, others argued that a statistical approach could not account for an unexpected and dramatic acceleration in achievement when a school, or group of schools, did particularly well. Disadvantages of this approach were also raised on the blog, for example the inappropriateness of such an approach when it is known that many of the weakest cohort no longer take the exam. Those on this side of the argument held that judgement was at the heart of the awarding process. Their view was that we need the right people from the subject community to be making those judgements. It was argued that the risk of taking the statistical route is that a student's result is divorced from the content of what they have learnt and the performance they have exhibited.

The consensus was that whilst this debate between judgement and statistics is a fairly technical one, it is also fundamental and a way needs to be found for an open and intelligible public debate about the move away from judgement to a more statistical approach.

Another view urged the need to tackle the various views that the public have about qualifications – that they should be differentiating, that they should be fair, that they should tell us what people are going to be able to do in the future – by getting them to think about those different claims and to try to reconcile them.

### Recommendation 3

There needs to be an open and intelligible public debate about the advantages and disadvantages of a move away from judgement to a more statistical approach, as well as a discussion about what role the public really wants exams to perform in the future.



## 04 Who should own qualifications?

Many participants in the debate felt the ownership of qualifications should be a partnership between schools, higher education, employers and awarding bodies.

Tim Oates argued that whilst awarding bodies had particular expertise in terms of measurement, they had to work in partnership with subject communities and users of qualifications who held legitimate views about the context and role of examinations.

One speaker noted the split that had occurred between the 'users' and the 'producers' of qualifications. Simon Lebus argued that since the creation of the National Curriculum, the British state has taken an ever-increasing role in mediating between subject communities, HE, professional societies, employers, teachers and examination designers. In other words, regulatory arrangements and politics have been inserted between users and producers. This has resulted in a divorce between users and producers, which has led to two very different views about what is happening in standards. It was argued that government should stand aside to allow an unmediated discussion between HE, employers, subject specialists, and awarding bodies, and that regulatory arrangements should be focused on weights and measures.

Others also questioned the role of the state, with some participants questioning whether standards could be unbiased and impartial when governments were transient and lacked necessary education expertise. Amanda Spielman argued that the state only has a role in setting minimum content standards. David Howe went a step further arguing the state shouldn't even do this. Instead its job was to facilitate a professional dialogue between the worlds of education and employment in terms of meeting needs, and ensuring that progress is made.

However, others gave reasons why there should be a role for government. Questions were asked about who would make decisions about qualification reform if politicians were excluded, given the number of people with an interest in the issue. Martin Rowland argued that government is the only medium that had the capacity and ability to bring together interested stakeholders to decide the purpose of qualifications. Gordon Stobart pointed out that if a government is funding much of what goes on in schools, it has an interest in the outcome in terms of, for example, international league tables. The question of where to intervene was therefore a difficult one. Isabel Nisbet referred to the need for a regulatory role and the importance of giving students some kind of assurance about the currency of qualifications.

Tessa Stone referred to the unintended consequences that occur when the relationship between businesses, HE and exam creation systems is in the wrong order. Concerns were expressed about HE and business beginning to create their own tests, even though they might not be experts in doing this. It was pointed out that university admissions tests were being created by those who may be subject experts, but are not experts in 'fair testing' for that sort of outcome. Similarly, with the debate around access to the professions, she argued that whilst the professions are being brought in to give their views about how to address inequities in the current system, they may not always be best placed to implement some of things they are intending, or to put this in the context of the wider education debate.

### Recommendation 4

Qualifications should be owned primarily through partnerships between schools, higher education, employers and awarding bodies. Regulatory arrangements should be focused on a 'weights and measures' agenda, allowing this unmediated discussion between schools, HE, employers and schools to take place. The State's role in relation to value for money needs to be achieved elsewhere in the education system.



## 05 How can employers be engaged?

The debate on ownership of qualifications led naturally to the issue of engagement of employers, and the mechanisms that are available for them to be involved in the qualification development process.

### Concerns from employers

Concerns were raised about the standards in key skills such as numeracy and literacy of those young people applying for apprenticeships. It was argued that those taking the vocational route are often treated as if they are on a lower track. They therefore come to Further Education colleges with very low grades and without the key skills needed as part of the apprenticeship. Others raised concerns that vocational qualifications were being seen as a way to ratchet up achievement by encouraging students to take courses which were GCSE equivalents but did not in fact contain a great deal of vocational work. The concern was that this was depriving children of a solid foundation in the mistaken belief that it is better to achieve a higher grade in something not useful than a lower grade in something more appropriate to their needs. This undermined 'real' vocational work.

### Ideas about future engagement

The audience reached a consensus that greater engagement with industry was necessary. One audience member referred to the work of the School of Communication Arts, a new institution which will deliver vocational qualifications. It is developing its curriculum with industry experts, using an online consultative process.

Many audience members felt we need a qualification system which is responsive to the end user. It was pointed out that where there is a degree of self-organisation amongst employers, there is a link between qualifications and work processes, and a link between qualifications and license to practise. Several were of the view that it should not be the case that we create a system which end users then have to find a way of adapting to.

### Recommendation 5

Our qualifications system needs to be responsive to the end user, rather than end users having to adapt to a system in which they have not had a part.



## 06 How important should qualifications be and how much can they be relied upon?

A completely different topic was the extent to which exams and qualifications should matter. Roger Murphy argued that despite the strengths of the UK examination system, *“UK examination grades are, and can only ever be, approximate indicators of student achievement”*. He referred to the simplicity of the grading system; the fact that any examination is a sample and that luck can come into play; the possibility of the same candidate performing differently on different days; and the use of judgements by those marking the exams, even in the most professional system.

Others argued that the evaluation of public examinations was in fact a highly technical area; that it was necessary to define and clarify terms to decide what kind of standards should be set and maintained; and what kind of evidence was relevant for setting and maintaining them.

In response to these comments about the need for a ‘scientific approach’ to educational assessment, Roger Murphy argued that whilst we should make UK public examinations as fair and meaningful as possible, we should accept that grade comparisons will always be approximate, and that examination results should therefore only be treated as one piece of evidence. His view was that whilst people should carry out specific educational assessments carefully and systematically, there needs to be a realisation that the topic is complex and not particularly amenable to an ‘exact science’ approach.

However, others questioned the notion that the same student would get a different grade if they did the exam on a different day. It was argued that in the upper ability range, students master the whole syllabus, and are therefore very likely to be able to tackle any question to a good level.

Nonetheless, this issue of robustness then served to raise a number of other questions. Some argued that if we are to rely on exam results to the extent we do at the moment, there needed to be a certain amount of commonality of achievement. In the House of Commons seminar, Dr. Robert Coe, from the

CEM Centre at Durham University, looked at the importance of comparability of grading standards for Higher Education, in terms of both selecting students for courses, and interpreting A level grades as indicators of future performance. He suggested that this becomes difficult if there are differences in standards over time, between subjects, and between examination boards, although he did concede that there is no way in principle to make A levels equivalent for all purposes.

Some suggested that there should be a debate about what else, apart from exams, should be in the mix for university admissions. Others suggested that the issue was with the artificial values that we place on standards and that something could be learnt from the Duke of Edinburgh Scheme, where everyone attains differently but everyone leaves with the same award.

### Recommendation 6

A public debate should take place to look at what alternative sources of evidence should be used for entry into Higher Education rather than simply relying on exams.



## 07 What are qualifications for?

There was a widespread view amongst the panel and audience alike that, in order to tackle the standards debate, it was fundamental to think about what qualifications are supposed to be for. The consensus was that the real debate was about the fitness for purpose of our current and future qualifications, rather than the focus on standards over time, which the media often focused upon. Other comments on the blog pointed out that determining the purpose that we want qualifications to fulfil in the future lay “*at the heart of the debate on standards*”.

Several participants referred to the multiple purposes that qualifications were now often expected to take on: accountability, driving up standards in schools, individual selection, and allowing greater access to education. It was pointed out that compared to 1975, we now have 45 per cent rather than 14 per cent of young people going into HE and, whilst the average student finished school at 15–16 in 1975, he or she now stays on until A level.

However, given this multitude of purposes, the view was that exams are not able to bear the weight of some of the purposes that are being put on them. Some suggested that in the case of A level, we need to look at its selective purpose as its primary purpose. In that sense, its fitness for purpose can be questioned because the increase in students with three As has made it more difficult to distinguish between them.

Others talked about the dichotomy between the employer side and the education side, arguing that HE wants something different to employers. It was argued that unless we decide what fitness of purpose is, there is no real point in talking about standards because industry employers want one thing and educationalists want another.

### Recommendation 7

Qualifications are often required to measure too many different things. To tackle this, there needs to be a public debate about what qualifications are for, with the possibility that we consider reducing the number of purposes for each qualification, focusing on a primary purpose for each of them.



## 08 What is the impact of the current accountability system?

Leading on from these concerns that the same instruments were being used to measure too many different things, many in the audience wanted to discuss what they saw as the distorting effects on education of a focus on grades. There was growing concern over what was seen as a 'performativity' rather than a learning agenda.

A number of participants argued that the problem was not exams themselves, but rather the way that the pursuit of high grades is affecting the quality of education. In other words it is not the results as such but the uses to which the results are put that is the problem. There was a view that there is currently far too great an interest in the headline results, the grades, and not enough transparency about the detail. Anastasia de Waal argued that because exams are used so often as proof of success, the focus has been pushed onto how good schools are at preparing children for exams, and the aim has become to ratchet up performance, rather than ratchet up learning.

There were concerns raised that the role of assessment has been transformed from fair and accurate assessment of what people can do to an agenda of narrow drilling towards exams. Concerns were also raised on the online blog about so-called 'performativity' incentives in the shape of league tables and associated target setting, and the impact that this was having on both examinations and the behaviour of teachers in 'teaching to the test'.

Specific issues around the focus on borderline C/D candidates were referred to, including the educational implications in schools, with neglect of the least able and most able. There was also a particular concern about the distorting effect of equivalencies (of other qualifications with GCSE), which was seen to allow the system to push weaker pupils into achieving higher grades but not by maximising learning. Some talked about the possibility of moving away from a tight system of equivalencies, allowing for a more flexible system where the currency value of the qualification itself will do the job.

Several people concluded that whilst the exams can be tweaked, it is the accountability agenda which should be the prime target of reform. Gordon Stobart talked about the pressure that exams are under in terms of the use of results in the accountability system and called England "*one of the fiercest accountability systems on the planet.*" It was argued that there should be a concentration on separating out what is needed for national accountability from what is needed for the individual student to be best prepared for their future. It was also argued that to do this, there should be less opportunity for government to intervene in terms of the pressure that is applied through the accountability system.

It was also concluded that we need to make sure we are looking at standards from a wider perspective so that the focus is not just on exam results, and that qualifications are about stimulating students and having a wider learning experience, not simply about exam preparation.

There was some disagreement about the impact that this emphasis on grades was having on A level students' choice of subjects. One audience member thought A level students used the data on pass-rates to inform their decisions about which A levels to take, picking subjects where there appears to be a better chance of gaining a higher grade. However, it was noted that Cambridge Assessment research had shown that A level students make their subject choices for reasons connected with employability.

There were also views about the distorting effect of assessment on the curriculum. Roger Murphy referred to the link between the taught curriculum and school exams in the UK as being a unique characteristic that others envy. However, he concluded that this was both a strength and part of the problem because when the curriculum changes there have to be matching changes in assessments. He believed assessment should follow curriculum, not lead it, with the focus on improving the standards of education and making education more relevant to young people and adults.

### Recommendation 8

As one way of avoiding the same instruments having too many purposes, there should be a concentration on separating out what is needed for national accountability from what is needed for individual students in order to ensure they are best prepared for what **they** want to do in the future.



## 09 Do international comparisons help?

Several members of the panel suggested that international comparisons might be helpful in terms of teaching and learning. For example, Roger Murphy said that we should be looking at whether other countries have found better ways of teaching than we have, rather than spending time focusing on exam results.

International comparisons were used as an example of how a greater engagement with employers can be achieved. The example of Germany was used to demonstrate the tradition of adapting the requirements of each industrial sector so that the timeframe and modification of the content of qualifications is entirely in tune with the sector.

Others suggested that international comparisons are useful because they are based on tests which students are not 'prepared for' in the same way that, as a result of the pressure of national league tables, they are prepared for examinations such as GCSEs and A levels.

### Recommendation 9

International comparisons are useful sources of information but care should be taken in drawing direct comparisons from them.



## Summary of recommendations

1. Before any discussion about 'standards', terms need to be defined and clarity reached about what kind of standards are being referred to. In particular, given the widely recognised problems with comparing standards over time, the assessment community questions the value of concentrating on this simplistic argument.
2. The frequency and scope of change in qualifications needs to be reduced, inappropriate change in the format and content of qualifications should be resisted, and the change agenda should be driven by the user and subject communities.
3. There needs to be an open and intelligible public debate about the advantages and disadvantages of a move away from judgement to a more statistical approach, as well as a discussion about what role the public really wants exams to perform in the future.
4. Qualifications should be owned primarily through partnerships between schools, higher education, employers and awarding bodies. Regulatory arrangements should be focused on a 'weights and measures' agenda, allowing this unmediated discussion between schools, HE, employers and schools to take place. The State's role in relation to value for money needs to be achieved elsewhere in the education system.
5. Our qualifications system needs to be responsive to the end user, rather than end users having to adapt to a system in which they have not had a part.
6. A public debate should take place to look at what alternative sources of evidence should be used for entry into Higher Education rather than simply relying on exams.
7. Qualifications are often required to measure too many different things. To tackle this, there needs to be a public debate about what qualifications are for, with the possibility that we consider reducing the number of purposes for each qualification, focusing on a primary purpose for each of them.
8. As one way of avoiding the same instruments having too many purposes, there should be a concentration on separating out what is needed for national accountability from what is needed for individual students in order to ensure they are best prepared for what they want to do in the future.
9. International comparisons are useful sources of information but care should be taken in drawing direct comparisons from them.



## Appendix 1: List of unnecessary changes and/or changes which threaten standards

1. Calculators in and out of GCSE maths.
2. Modularisation across the whole system, without piloting, in 2000 (awards in 2002).
3. Wholesale modularisation in GCSE in 2008 (awards in 2010).
4. The drive towards reduction in the gross number of qualifications in the English system, plus the use of the concept of 'coherence' – overturned operationally by the decision to allow FE centres awarding powers.
5. Overblown content of the National Curriculum followed by extreme, inappropriate reductions (e.g. KS4 Science) – this impacts on awarding bodies principally due to GCSE requirements to cover National Curriculum.
6. Core skills and functional skills as a hurdle.
7. Languages in and out of the curriculum requirement.
8. Shifts to and from coursework, and in the form of coursework; 2002 GCSE criteria from QCA determining that all GCSEs should include coursework, contributed to bringing coursework into disrepute by invalid/unpopular inclusion and increase in assessment volumes.
9. Different messages regarding equivalences and changes in equivalences.
10. Reduction from 6 units to 4 (again wholesale system change rather than per subject); awarding bodies proposals for 4-unit qualifications over-ridden in 2000.
11. Reduction in assessment time from 4 hours to 3 hours in 2002.
12. The pace of the accreditation cycle, leading to changes after one year of a specification. This highlights that change is now five years rather than 10.
13. The fight over the form of stretch and challenge and A\*.
14. Constant change in the name, but not the form and content of, core skills, key skills, functional skills.
15. GNVQs, AVCEs, Diplomas. No stability in the vocational route. Constant academic drift in successive transformations of the vocational route.
16. In national assessment, mental maths.
17. Split and merger of English Literature and English Language and current perverse performance table rules.
18. Unreasonable subject changes  
e.g. <http://www.tes.co.uk/article.aspx?storycode=337286>
19. Maths tiering – access to grade C in tiered examinations.
20. Fluctuations in the assessment of SPG in GCSEs.
21. Vacillations on Diplomas=only qualification versus Diplomas=distinctive route for vocational.
22. AS standards-rating fixed at 50 per cent of AS+A2.
23. Decision by QCA to combine maths and further maths – subsequently overturned.
24. Detail in Languages – in February 2007 it was agreed that a pilot of Oral Language Modifiers (OLMs) would be held in summer 2007. There were three salient features; (a) the use of a sample of students, (b) it should establish whether the use of OLMs compromised the integrity of the assessment, (c) it would establish how a wide range of disabled students would operate with OLMs. Five weeks after this agreement letters were written changing all three features; (a) there would be no sampling but all deaf candidates would be eligible, (b) it would not inform the impact of OLMs on the integrity of the assessment but would simply inform good practice, (c) no students with disabilities other than deafness would be included. Awarding bodies thus found themselves committed to what was effectively a different project upon which they had not been consulted.

## Appendix 2: Examples of past examiners' reports

### **UCLES Examiners' Reports, 1858:**

"[School leavers demonstrated] little indication of an acquaintance with the best elementary mathematics works."

### **HM Inspector's Report, 1876:**

"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."

### **UCLES Examiners' Reports, 1920s:**

"Punctuation was almost universally deficient or valueless."

### **Examiner's Report, Pure and Applied Maths, 1924:**

"The only point that calls for a report is the general weakness of a large proportion of the candidates."

### **Examiner's Report, 1932:**

"A considerable percentage of the candidates were quite unfitted to take the exam and had no possible chance of passing."

### **Report of the Consultative Committee on Examinations in Secondary Schools, 1911:**

"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."



CAMBRIDGE ASSESSMENT

Cambridge Assessment  
1 Hills Road  
Cambridge CB1 2EU  
United Kingdom  
tel +44 (0) 1223 553311  
fax +44 (0) 1223 460278  
[www.cambridgeassessment.org.uk](http://www.cambridgeassessment.org.uk)



UNIVERSITY of CAMBRIDGE  
Local Examinations Syndicate

Cambridge Assessment is the brand name of the University of Cambridge  
Local Examinations Syndicate, a department of the University of Cambridge.  
Cambridge Assessment is a not-for-profit organisation.



Cover: Corbis Images