

**Higher Education Policy Institute (HEPI) and
Joint Information Systems Committee (JISC)**

House of Commons Seminar: 27th January, 2010

Quality and Standards

Bahram Bekhradnia (Chair):

Ladies and gentlemen, I welcome you to the first event in our House of Commons seminar series. We're very pleased to be partnering with JISC, and we are here through the good offices of Evan Harris MP.

The topic for today is "Quality and Standards", an issue that's never really been absent from the debate and has been particularly prominent recently. Last year, for example, the former Chief Executive of the QAA described the degree classification system as "rotten". That was misinterpreted by much of the media as a comment on the quality and standards of the whole higher education system, which of course it was not. The Select Committee, in a report last year, commented extensively on the question of standards without necessarily displaying terrific understanding of the issues – I have to say, having sat in myself on the discussions between the Select Committee and the group of Vice-Chancellors, it was like the deaf talking to the deaf. And, of course, in HEPI we have produced data on the amount of private study that students do - again conveniently misinterpreted as being reports on contact hours and issues of pedagogy.

There are a host of issues and there's a lot of cloudy thinking around them. I hope that some of that may be exposed today. We have a distinguished pair of speakers. Peter Williams, the former Chief Executive of the QAA, has never hesitated in saying what he thinks (even when he was in office) but we hope that today he'll say what he *really* thinks. And most of you will know Roger Brown as one of the country's leading theoreticians on the questions of quality and standards; he has written and spoken extensively on the subject.

Peter Williams:

Quality and standards in higher education – it's a very vexed and tricky topic and one that isn't going to go away. It's one in which everybody's got a view, but rarely are those views informed by fact. I want to spend quite a lot of the short time available establishing some facts.

The subject is particularly tricky for a number of reasons. Perhaps the most important of these, and a good starting point, is the lack of any generally agreed definition of what quality is and standards are. So what is 'quality' and what are 'standards' in the world of higher education? We all use the same words, but we mean different things by them. To avoid talking at cross purposes, however *you* may use the words, this morning *I* am going to use them like this.

When I talk about good quality of higher education, I shall mean courses which have a clearly defined purpose, which are specifically designed to meet that purpose, which conform to the requirements of the National Qualifications Framework for Higher Education, which are delivered using pedagogical techniques that are fit for purpose and regularly checked for effectiveness, and in which students are assessed against national expectations using

assessment methods that will measure reliably what they are intended to measure. I think it's very important that we don't just use 'quality' in a general way if we're going to expect it to have specific significance and specific outcomes.

When I say 'standards', I am going to mean explicit and predetermined levels of knowledge, understanding and skill that are required to be demonstrated before a degree is awarded, and which meet the expectations of the National Academic Infrastructure. (If you don't know what the National Academic Infrastructure is, then you ought to, and you'd better see me afterwards.)

I shan't be any more prescriptive than that. Basically, those definitions mean that higher education should do what it says on the tin. That means there has to be a tin, it has to have something in it, and it has to have a description of its contents on the outside. The big question at this troubling moment that we're living through is – who should decide on the contents of the tin? Is it the producer or the user?

Now, onto those facts that I said I would try to pin down. They are mostly about standards - I'll talk about quality a bit later on if time permits. There are six bodies in the United Kingdom that award A Levels (and other secondary education qualifications, such as Highers in Scotland) – one based in Scotland, one based in Wales, one based in Northern Ireland and three based in England – and you know how difficult it is every year to get anyone to accept that the standards are comparable. There are always arguments about which is the easier board.

In contrast, there are about 156 higher education institutions and they are all legally entitled to award degrees in the United Kingdom. Together, they offer nearly 50,000 identifiable degree programmes. Universities award their degrees independently as autonomous organisations. There is no national assessment system, the degrees belong to the awarding institutions severally, not collectively, and they are not owned by the state as they are in most European countries. Assessment of students is entirely the responsibility of the awarding institution. In practice, this means the relevant department is overseen by educational faculty, academic administrative officers and a variety of committees.

In an attempt to ensure that their degree standards – that is, what they are demanding of their students – and the fairness of their assessment procedures meet the collective expectation of the national higher education community and other stakeholders, they employ examiners from other institutions. It isn't known how many external examiners there are, but it's probably upward of 20,000. External examiners are contracted by the awarding institution, which tells them what duties they are to perform. In recent years the QAA has provided guidance on the minimum expectations of an external examiner but, crucially importantly, there is no nationally owned or controlled system of external examiners. Many of them examine students on more than one programme.

Almost all Honours degrees are classified as being either First Class, Second Class Upper Division, Second Class Lower Division and Third Class, and passed non-Honours degrees can also be awarded. There is no common definition of these classes, though there are a number of generally similar locally used definitions. The meaning of the term 'Honours' itself is no longer clear as there are relatively few non-Honours degrees on offer, except in Scotland where they still form a significant proportion of the degrees awarded. And, by way

of further enlightenment, I'm indebted to the Student Room website for the following insights into degree classification:

A First is known as a Geoff or Damian Hurst;
A 2:1 is known as an Attila the Hun;
A 2:2 is known as a Desmond Tutu; and
a Third is known as a Douglas Hurd.

In addition, 2:2s are often light-heartedly referred to as a drinker's degree with the implication that the graduate spent more time in the Students Union bar than studying. A Third is sometimes known as a Richard (after the monarch) or a Vorderman (as British TV celebrity mathematician Carol Vorderman only received a Third at university). Finally, a Pass degree is sometimes known as a Khyber.

Degree certificates generally offer no information about a student's achievement over the course of their studies other than the name of the degree, its subject and the class and date of the award. Many universities, but not all, also offer a more detailed transcript providing information about the courses taken and the marks obtained. Since 2005, as part of the Government's undertaking, without any reference to the institutions, that the UK's higher education system should conform to the norms of the Bologna process, there has been an expectation that all graduates should receive free and automatically a transcript that complies with the rules of the European Diploma Supplement, a common form transcript devised a decade or so ago. At the last count, bearing in mind this was supposed to have been done in 2005, only about 65% of institutions were issuing EDS compliant transcripts with a very much smaller proportion doing so free and automatically. The UK Europe Unit has been undertaking an inquiry into the prevalence of the European Diploma Supplement, so we expect further information to appear about this soon.

There are many other facts I could regale you with, but the point I want to make is that the current means of assessing and classifying students is essentially local rather than national, notwithstanding the external examiner system. I believe there is a broad comparability at the basic threshold degree level, which is manifested essentially within the system and not very easy to get at from outside, but beyond that the heterogeneity of the system becomes more relevant than its homogeneity.

So it continues to amaze me that desperate and, to my mind, futile efforts are made to pretend that all Firsts are the same and all Upper Seconds are the same. I don't think that this pretence is in anybody's interest. The less prestigious institutions suffer from unequal competition with the more prestigious and are not able to play to the strengths of their diversity, while the more prestigious find themselves in wasteful competition with their peers, using doubtful statistical legerdemain to ensure that they are at the front of the pack. The students - and in a sense it's the students that are the biggest victims of this system - find themselves in a system which views a 2:2, a respectable pass, as, in effect, a failure.

There is no mechanism to ensure consistent and meaningful comparability among institutions and subjects, and no mechanism I can envisage that could make it so. National examinations, which some have suggested, or individual degree standards overseen by a body such as QAA, would create a vast industry and an attendant bureaucracy and its inevitable failure would make the annual row over GCSEs and A Levels look very tame indeed. It would be much simpler to stop using these out of date classifications designed to meet the needs of another

century, and provide individually focussed information which actually tells the user something about the student and what he or she has learned. The 'one size fits all' scheme we now use is a travesty of fairness and consistency. After all, medical degrees aren't classified, and it doesn't seem to matter that your doctor can't say whether he or she was First Class or Third Class. I do have great hopes that the report into higher education achievement that Bob Burgess is pursuing will come up with some answers to this problem.

But the question of quality and standards isn't just about degree classes. We seem in this country to have no capacity to think beyond monolithic hierarchies and, in trying to shoehorn very different purposes, clienteles, structures and people into a single narrow boot marked 'The only acceptable HE standards for the UK', we perhaps reduce our opportunities to innovate, develop and recognise a much more useful set of standards based on the particular characteristics of the students and programmes being offered. Why should the standards needed to gain a degree in Classics be the same as those required by a degree in Medical Science or Social Work? Each subject has its own aims and objectives, its own purposes and methods, its own expectations and requirements, and I believe that these should be the basis for standards. Provided the standards are clearly stated and readily available, validated by the relevant subject and professional community as useful, valuable and appropriate, and form the basis for the assessment of students, then the variations between subjects and institutions should become a reason for celebration, not the sort of angst about irreconcilable differences.

And what of quality? I defined quality at the beginning, and said that it was different from standards. Is quality declining? Does it matter? Well, I believe that quality should be assessed in relation to what it says on the tin. Teaching and learning support systems come in all shapes and sizes, some of them unrecognisable from when I was a student. So long as what is provided is consciously designed to ensure that the student meets the intentions and objectives of the study programme on offer and can be shown to do so through independent assessment, then I do not believe there is a case for the kinds of complaints about these or other perceived shortcomings. Value for money is not, I would suggest, to be found in the number of hours of face-to-face contact offered - lectures where nobody learns anything because those attending are too busy tweeting or texting, or seminars where no-one's actually read the text, or labs where the students are given the same problem to solve two weeks running - no. The value exists in the professional care taken to devise a learning scheme that exposes the student to a variety of educational stimuli, making sure that they can use these to learn the things they need to know while, at the same time, developing other skills or routes to understanding. This may involve little direct contact with staff, but a great amount of engagement with learning. Using that test, which is not recognised by the traditional bean counters, there are many institutions and academics, though not enough, who are truly involved in the education of their students, not just in preparation for examinations or training for a narrow career horizon.

So what, if anything needs to be done about quality and standards? I have seven things I would suggest:

1. Scrap degree classes - they're a dangerous diversion from the real business of higher education;
2. Require degree courses to be described explicitly, but allow them to be described in their own terms. Make sure that the descriptions of standards are readily

available to potential students and employers and sufficiently detailed to be readily tested by them;

3. Provide real information about graduates' achievements;
4. Ensure that external examiners all work to at least a common minimum job description;
5. Judge quality in terms of its fitness to deliver the programme's aims and objectives, not against notional quantitative norms;
6. Welcome pedagogical innovation so long as its intention is to maximise learning and personal development and not just to cut corners and costs;
7. And from time to time, say about every ten years, give each institution a more thorough examination than they get at the moment through audits and certify them as fit for public recognition.

I put those on the table and leave it to Roger to demolish them.

Professor Roger Brown:

Thanks very much. I always try to start with an appropriate text for what I'm going to say. I don't know how many people remember Peter Cook and his interviews; one of his interviewees was a man called Sir Arthur Streeb-Greebling who had spent his entire life trying to teach ravens to swim under water, but not so far successfully. And his famous comment was, "I've learned from my mistakes. I'm sure I can repeat them."

However, I'd also like to present three slightly more serious quotes. The first comes from Peter's estimable Agency in 2006: "It cannot be assumed that students graduating with the same classified degree from different institutions, having studied different subjects, will have achieved similar academic standards; that students graduating with the same classified degree from a particular institution, having studied different subjects will have achieved similar academic standards; or that students graduating with the same classified degree from different institutions having studied the same subject will have achieved similar academic standards."

My second quote comes from the Select Committee report that Bahram referred to: "We have concluded that the QAA should be responsible for maintaining consistent national standards."

My third quote is from the HEFCE consultative document of December 2009: "Revised QAA arrangements will need to address (a) concerns that standards between institutions are not comparable or consistently applied."

So what I want to cover in my 15 minutes is, first of all, what does comparability actually mean? Why has it become an issue? Is comparability even possible, and, if it isn't, what do we do about it? I will try to deal with a complex subject as simply as I can.

I have the same basic definition of standards as Peter does - levels of academic achievement necessary to attain a specific university award. In other words, the levels of knowledge aimed at and achieved following a specific programme of study and indicated by the award

and, as I think most of us know, standards are set when programmes are designed and validated, but they are only finally fixed when students are assessed.

Comparability means that the standards of learning aimed at and achieved by students in any two programmes leading to the same or a cognate award are genuinely equivalent. So it could mean, for example, that all students in one institution obtaining a bachelors degree in any subject are achieving the same standard, all students from several institutions obtaining a bachelors degree in any subject are achieving the same standard, and it could mean all students from several institutions obtaining a bachelors degree in the same subject are achieving the same standard. It could also refer to common standards in all elements of a programme, options as well as core, and it could mean common standards over time in different cohorts of a programme. For the moment, until I demolish it, I'm going to talk about the hard meaning. In other words, that all students are achieving genuinely equivalent standards of achievement.

Why has comparability become an issue? Well, again I have a couple of quotes here. The first is from the Council for National Academic Awards' Second Policy Statement April 1965: "Since the Council was established with the purpose of enabling colleges to plan their own courses and to admit and examine their own students, it will impose any such basic requirements as are necessary to ensure that its degrees are comparable in standards to those of the universities." The CNAA did that, through, in its early days, mainly using academics from old universities.

The other quote comes from the mid-1980s from the first Universities UK Code of Practice on External Examining: "The purposes of the external examining system are to ensure first and most important that degrees awarded in similar subjects are comparable in standard in different Universities in the United Kingdom." Once again the mechanism was peer review, through, in this case, external examiners.

I don't need to go on about the benefits of comparability. If comparability could be achieved, it would mean that everyone could rely upon the currency of awards, students, employers, funds, etc. Internationally, the currency of the British degree depends in part upon a perception of common standards. However, we have to understand that comparability is under great pressure. There are three particular pressures. First of all, we have longstanding and depressing evidence of weaknesses in assessment – I found the first reference to weaknesses in university assessment dates back to 1888. There is a very authoritative article in University Court in 1967 particularly picking on reliability. Here is a quote from the QAA 2008: "Worries include doubts in some cases about the double marking and/or moderation of students' summative assessment, continuing difficulties with degree classification, departures from institutional practice in the way staff in departments and schools work with external examiners, and generally weak use of statistical data to monitor and quality assure the assessment of all students and degree classifications." So there is a longstanding problem that we have as a sector in assessing students in a professional fashion.

The second pressure, of course, is the enormous expansion and diversification in the system – a huge increase in the numbers and types of students and subjects, an increase in the categories of work being assessed (not just invigilated exams but all manner of other things), the growth of multidisciplinary and modular programmes and the increasing importance of things like transferable skills, enterprise and employability that are very difficult to assess by conventional assessment means.

The third set of pressures, which were partly what motivated the Select Committee, were the increasing competitive pressures on institutions, including evidence in a small number of cases of management intervention with academic examiners' judgments. As we go into a resourcing squeeze, we need to bear in mind that in a commercial market lowering quality whilst maintaining price is a legitimate response to competitive pressures.

You can't divorce all this from what's going on in the system more widely. I would argue that these pressures on comparability are an inevitable consequence and corollary of national policies to expand the system, to diversify provision, to increase efficiency, to make the curriculum more responsive to the economy and, particularly, to enlarge student choice. At the same time resources have been under pressure, research has continued to have priority over pedagogy in many institutions and, of course, market competition has become much more important.

I'd like to refer to what I've called Brown's Paradox (but I don't claim originality for it) which is that, as the system expands, the pressures of comparability increase but, by the same token, the ability to ensure it reduces. From one of the HEQC reports that went into the Graduate Standards Programme – this was published 12 years ago: “Overall, it is felt that assessment in higher education is not compatible with a threshold model in the sense of a sector-wide set of outcomes that will define explicitly and meaningfully what either a particular sort of graduate or any sort of graduate actually is.” Indeed the major changes that have taken place over the last decade have produced an incredibly heterogeneous sector with far more types and structure of degree than in the past. And this looks set to continue. They make such threshold standards increasingly impossible to implement, at the same time as creating a situation which makes their absence felt, and I think that is the nub of the problem.

Onto my third question. Comparability is certainly desirable, but is it actually feasible? I would argue, a bit like Peter, that if you really want to have genuine equivalents in standards then a number of quite demanding conditions need to be met. The academic achievements of the students would need to be genuinely comparable, and that would require entering students to have or be capable of acquiring initially comparable knowledge. It would require comparable course outcomes and probably common course content. It would require common marking practices in each component of each course and common assessment tasks, and common rules about the way in which marks are combined to produce a degree classification or some other summative outcome. But, as I'm sure people who have studied the subject will agree, if you look at the literature on assessment in the UK, you will find not only considerable variations between subjects between institutions, but within subjects between institutions. Some of these variations seem very hard to justify; others are inevitable because of the different forms of knowledge production found in the various academic disciplines, and because of the need for examiners to have discretion.

But assuming all that could be taken care of, is strong comparability really desirable? Should a demonstrable persistently significant lack of comparability mean some exam boards, departments or even possibly institutions giving larger numbers of highly rated awards and others fewer? Would some courses have to teach less or to a lower standard and vice versa? Should there be changes in resourcing levels and policies in ambitions, criteria, etc? A combination of some or all of these might put certain programmes, departments or even, dare I say, institutions, out of business. Who would decide these things assuming we were to get that far?

I believe that any real comparability now is infeasible, at least without a national curriculum and national examiners answerable to a national standards agency. With respect to Peter's example, even with only six exam boards, almost every month in the Times Educational Supplement there is a discussion about comparability of standards in the schools. For example, in March 2009 the QCA found that nearly half of the 14 year olds' English tests were wrongly marked. So I'm afraid I think that comparability in a strong sense, even if it may be desirable, is quite infeasible and would not be a good use of resources if we were to attempt it. But that doesn't mean we can simply *laissez-faire*.

So here, finally, would be my responses, and they're not a million miles away from where Peter has got to. First of all, we have to be much clearer as a sector what we mean when we talk about standards and comparability. For example, HEFCE's consultative document put out before Christmas is frankly muddled; you can read several different meanings of comparability into it and HEFCE needs to clarify what it means. Secondly, I completely agree with Peter that institutions need to be much clearer about what is on the tin and what it is that the programmes that they are offering are equipping students to achieve - programme specifications are obviously a mechanism for that. Thirdly, I think we have to re-visit the subject benchmark statements to see whether more can be done with those without, I have to emphasise, fettering institutional responsibility or examiner discretion.

I think my most important suggestion is that we need to pick up where HEQC left off. We have to create subject-specific networks that enable staff to develop stronger shared understandings about standards by comparing course aims, mark schemes, etc. That was, in fact, the central recommendation of the Graduate Standards Programme. Let me read you a brief extract from the Graduate Standards Programme Report of 1997: "We recommend that subject groups and professional networks should develop or extend opportunities to build common understandings and approaches among academic peer groups. For example, maintaining expert panels for validation, accreditation, external examining, assessing, developing directories of programmes, programme elements by field, relevant programme award progression statistics, statistics about progression into employment, mechanisms to monitor changes in standards at other educational or occupational levels and internationally, formal opportunities to discuss and review standards." The external examiners attempt to do this but, frankly, they can only do it to a limited extent.

The report also recommended more systematic training and development for institutional assessors and examiners, the creation of internal examiners forums, separate from exam boards, to review assessment practices and share perspectives, the development and use of assessment archive, data and information to help with monitoring standards, trends and awards over time, relationships between higher education and other wards, inter-institutional benchmarking of assessment regulations and protocols to compare and agree common conventions on regulations – threshold pass marks, the algorithm used to define the final standard of award, compensation rules. That, I think, is the most powerful thing that the sector could do.

We also need to improve our assessment practices. The QAA said something about that in its Thematic Enquiries Report last year, but it's singularly absent from the HEFCE consultative document. I can tell you if that were done, it would achieve far more than the rest of the HEFCE consultative document put together. Clearly we need to phase out degree classification, and I also think that the time has come to blow the whistle on external

examiners. I think we need a new system of departmental academic review, very different from teaching quality assessment and internally owned by institutions which would tackle particularly the question of the worthwhileness of the aims of programmes.

To sum up, my view is that real comparability in the sense of genuine equivalents of standards as defined is infeasible in a mass diverse system. I've indicated that there are a number of things that we can do to limit variability: in particular, if we are serious about comparability, we have to increase the amount of systematic benchmarking we do through assessor networks. We can't rely upon external examiners. Indeed, rather than ask them to pursue what can only be described as heroic tasks, we need a new system of academic review which would subsume the functions that they fulfil. Otherwise, I'm afraid we shall be having this seminar again in five years' time and Sir Arthur Streeb-Greebling's remarks will be even more pertinent.