



Implications of 'Dimensions of quality' in a market environment

Graham Gibbs



Contents

	Page
Foreword by Professor Craig Mahoney	4
Ministerial Foreword by Rt Hon David Willetts MP	6
Executive summary	8
1 Introduction	12
2 Definitions of 'quality' underpinning student choice	13
3 Implications for quality assurance	15
4 Implications for cost-effectiveness	18
5 Implications for quality enhancement	20
6 Reviewing teaching quality indicators annually	22
Case study 1	25
7 Relating programme-level quality indicators to collections of courses	26
8 Retreating from modularity	27
9 Focusing quality enhancement on programmes and course teams	29
Case study 2	30
10 Adopting new institutional change processes	32
Case study 3	33
Case study 4	35
11 Focusing on 'hygiene' factors and consistency of delivery of services	36
12 Changing students' role from consumer to active participant	37
13 Competing research agendas	38
14 An increased emphasis on institutional 'USPs' and distinctiveness	39
15 Rebuilding learning and teaching strategies	41
16 Final comments	43
17 Policy implications	44
References	47



Foreword

The Higher Education Academy (HEA) published *Dimensions of quality* in 2010, at a time when the perennial debate about what constitutes quality in higher education was reignited by the publication of the *Independent Review of Higher Education Funding and Student Finance* by Lord Browne. Even though the review focused on the English HE system, it heralded intense debate across the UK higher education sector and paved the way for a renewed focus on putting the student at the centre of the learning experience.

The debate about quality remains as pertinent and topical as ever. It is with great pleasure, therefore, that we publish this new report from Graham Gibbs that examines how the HE community can apply the dimensions of quality he identified in 2010 to make a real difference to the student experience.

Recent changes in the sector have brought with them a renewed focus on the provision of information to prospective students. The publication of the Key Information Sets (KIS), which include the results of the National Student Survey (NSS), is designed to provide prospective students with more information about the quality of higher education institutions throughout the UK. However, Graham, in this report, questions whether the information provided does tell students what they are likely to gain from a higher education experience at a particular institution. As a national body committed to enhancing the student learning experience, it is essential that the HEA understands, and supports others in understanding, what information students need to make the best of their time in higher education. This research will help us to do that.

The research also notes the disjuncture between a quality assurance system focused on the module and the NSS's provision of feedback on the course or programme. He challenges the sector to refocus enhancement strategies on the whole degree programme and on the development and recognition of the teaching team, which delivers it, rather than just on the module and the individual tutor as at present. The HEA's wide-

ranging programme of staff development and its accreditation of teacher development programmes via the UK Professional Standards Framework make it exceptionally well placed to respond to this challenge, and to lead change and innovation in this area.

Implications of 'Dimensions of quality' in a market environment is the latest piece in a programme of research commissioned and published by the HEA to provide the sector with an evidence base for policy-making in learning and teaching in higher education. We very much look forward to working with the higher education community to make the best use of this evidence base to improve learning and teaching, and to enhance the student learning experience.

A handwritten signature in black ink, reading "C Mahoney". The signature is fluid and cursive, with a large initial "C" and a long, sweeping tail on the "y".

Professor Craig Mahoney
Chief Executive, Higher Education Academy
November 2012



Foreword

I am grateful to Professor Gibbs and to the Higher Education Academy for producing this report. I welcome this valuable contribution to the debate and the challenge it provides. We share the commitment to the quality of the educational experience in higher education.

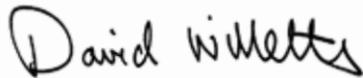
Higher education in this country has never represented a single experience or learning environment. One of the great strengths of our higher education sector is its innovation and diversity. With people's lives becoming increasingly busy and complex their aspirations for learning, whether starting a career or seeking to enhance one, are increasingly diverse. New technologies and even more flexible modes and routes of study are part of the story but whatever the route of study, at its heart is teaching. Teaching - the activities of educating or instructing, activities that impart knowledge or skill - should be at the heart of our higher education sector and the mission of every higher education institution.

Every student deserves an excellent higher education experience. No one would argue with the view that an individual's learning experience and success depends as much on their own engagement as with the actions of the institution and its staff. Effective learning is a two-way street. Yet it is often easier to find out about an institution's social activities than it is to find meaningful information on the teaching environment or the course outcomes.

Prospective students need access to a range of information so they can make effective choices on where and what to study. They will be more active and engaged learners if they understand the options available and choose the learning environment that suits their own needs and aspirations best. I am proud that this year the Government has been able to launch the Key Information Set (KIS). This is a first step. The KIS is not the conclusion of an agenda to improve information and student choice, though it is a very real step forward. We must evaluate and continue to

improve the KIS but not all the information that students deserve will fit a KIS model where data has to be produced on a comparable basis. As Professor Gibbs identifies, richer information can and should be delivered locally, in a prospectus or other course materials – or on websites that supplement the KIS data with bespoke information. His observations will contribute to the review of the KIS which will be carried out next year.

The power of this report is in the case studies. They provide an insight into how institutions are using indicators and measures of quality to engage actively with their students. They show how effective feedback and assessment improves and enhances the learning experience. I welcome Professor Gibbs' contribution to this important debate and encourage others to engage and add to our knowledge and evidence in this area.

A handwritten signature in black ink that reads "David Willetts". The signature is written in a cursive, slightly slanted style.

Rt Hon David Willetts MP
Minister of State for Universities and Science
November 2012

Executive summary

- 1 This report concerns the practical implications of the use of performance indicators for the way institutions are currently attempting to attract students, improve quality, improve 'value for money', and improve their relative standing in relation to educational provision. Institutions are responding to this data-driven market in a variety of ways, some of them perhaps unexpected and some with probably negative consequences. The report suggests ways in which the use of data in a market could be tuned up to have more positive effects.
- 2 The conclusions of the report are based on:
 - examination of the data currently available to students and used by institutions, and their validity and usefulness;
 - literature about performance indicators in higher education, and also literature about the effect that performance indicators and markets have on the behaviour of organisations in any public sector, such as schools and hospitals;
 - meetings with those senior managers responsible for educational quality within institutions, both in national gatherings and through interviews within 12 institutions of a wide variety of types;
 - examination of institutional documentation, for example about how quality data are reported and used internally, and institutional responses to the Browne Report.
- 3 It is not yet clear whether institutional attempts to improve National Student Survey (NSS) scores and other quality indicators is having any effect on student recruitment, let alone on learning gains. To a large extent the market is perceived to be driven by reputation, just as in the past. US research shows that reputation tells you almost nothing about educational quality, use of effective educational practices, or learning gains, but merely reflects research performance, resources and fee levels. It is uncertain whether the use of more valid indicators of educational quality will gradually change perceptions of what reputation is about, and turn it into a more useful guide to student choice.

- 4 Data currently provided to potential students, such as Key Information Sets (KIS), and used by institutions to make decisions, include some valid indicators of educational quality and also include variables that are invalid or difficult to interpret. There is scope to improve the value of the information provided to students, and used by institutions, by changing some of the variables and collecting and collating somewhat different data. In particular it is not yet possible for students to see what educational provision their fees will purchase (such as class size, which predicts learning gains) other than the proportion of class contact hours (which does not predict learning gains).
- 5 The aspects of educational provision that institutions pay attention to in their internal quality assurance processes often overlook crucial indicators. Any new quality regime should ensure that it focuses on the right variables, and the use of valid quality indicators in KIS and elsewhere would help to lever appropriate attention.
- 6 Regardless of the validity of currently available data, institutional behaviour is being driven by data to an unprecedented extent. In most institutions there is now an annual cycle of analysis of performance indicators at both institutional and departmental level, followed by planning to improve them, again at both institutional and departmental level. Departments are much more aware of how their competitors at other institutions perform, in relation to the main indicators. In some cases this annual analysis of data has in effect taken over from periodic review and QAA audit as the main driver of quality assurance and enhancement (and without this having been planned or agreed). Any future revision of national quality assurance mechanisms, and requirements on institutions, will need to take this reality into account.
- 7 Most currently available data are about degree programmes, and students apply to study degree programmes. In contrast much quality assurance, and course design and documentation, has focused on individual modules. In modular course structures the collection of modules that students experience may relate loosely to the unit of analysis of the NSS. This confronts modular institutions and modular degree programmes with major problems in interpreting and acting on the degree-programme-level data from the NSS. A consequence is that some institutions are greatly reducing the number of combined Honours degrees offered and moving away from modularity back to traditional single subject degree programmes with greater alignment of student experience with the unit of analysis, and labelling, of public indicators of quality. There are consequences of this shift for the diversity of curricula and for student choice, which may have negative impacts.

- 8 There has been a considerable emphasis over the past decade on training and accrediting individual teachers, rewarding individual teachers, and on funding local innovation in teaching. There is a marked lack of corresponding institutional emphasis on the effective operation of 'programme teams' (all those who contribute to the teaching of a degree programme), on developing leadership of teaching, and on curriculum design and assessment at programme level. A change of focus of national and institutional enhancement efforts is overdue. Institutional career structures still need to be developed that reward leadership of teaching, rather than only individual research and individual teaching. Funding for innovation, both within institutions and by national bodies, should be targeted on programmes rather than on modules and on the involvement of entire programme teams rather than on individuals.
- 9 Many institutions are using data to identify a previously overlooked quality problem and address it: the most common example is poor and slow feedback to students on their assignments. Institutions are then making very broad scale changes that affect all degree programmes and all teachers in order to address these problems. Data are successfully driving change and in some cases there is clear evidence of improvements in NSS scores as a consequence of the institution-wide change. Some centrally determined changes will limit teachers' scope to enhance teaching in contextually sensitive ways, and will make things worse.
- 10 An increasing number of institutions are using data to track progress in emphasising the 'institutional USP'. They are marketing themselves as distinctive in relation to a particular indicator, such as employability, and emphasising that variable in programme-level learning outcomes and in institution-wide quality enhancement efforts, and then collecting better data than are currently available in order to monitor progress.
- 11 In light of the prominence given to overall student satisfaction data in KIS and league tables, it is not surprising that institutions are addressing 'satisfaction' issues with vigour. This may be less to do with teaching than with consistently high standards of service delivery. In some cases these two domains of quality overlap, as with policies and practices concerning assignment turnaround times. Many institutions have a range of initiatives designed to improve service delivery, using NSS data to target efforts.

- 12 While there is a sense in which students are being treated as consumers of a product, institutions with good and improving NSS scores often have initiatives that engage students as co-producers of knowledge, or partners in an educational enterprise. Attempts to improve student engagement are taking many forms and sometimes involve students having responsibility for administering and interpreting student feedback questionnaires, and acting as change agents, and also central support for activities run by the students' union that relate to educational provision. It is unclear the extent to which NSS scores for a programme reflect extra-curricular initiatives of this kind, but some institutions are behaving as if they are important.
- 13 One focus of attention of the interviews undertaken for this report was whether institutions are focusing on 'value for money' by paying renewed attention to using cost-effective teaching methods in order to deliver a good quality of education given the level of fees and other income. There seems to be plenty of evidence of a squeeze on resources, and adoption of practices that save money, but not of an equivalent focus on using more effective methods. There is a need for a national initiative on cost-effective teaching so that, where reduced resources force changes to teaching practices, it might be possible to maintain or even to improve student learning.
- 14 Some of the institutions that are charging the lowest fees are suffering from competing demands to maintain or enhance their research efforts in order to retain research degree awarding powers. Attempts to improve teaching quality in such contexts face challenging conflicts of interest.

I. Introduction

Dimensions of quality (Gibbs, 2010) identified those educational variables that predict how much students learn. It was commissioned at a time when there were concerns about the quality of UK higher education relative to its international competitors, about the relative quality of very different UK higher education institutions, and about the adequacy of quality assurance regimes to assure that quality and prevent it varying too much between institutions. *Dimensions of quality* identified the quality indicators it would be valid to take into account when making such comparisons, and identified those indicators that, while they are often used, are invalid or difficult to interpret.

The agenda has since moved on to providing potential students with information about the educational provision their fees will purchase and using market forces to lever up quality. The National Student Survey (NSS), Key Information Sets (KIS), and league tables created in a variety of ways, now provide quantitative indicators to an extent not previously publically available. *Dimensions of quality* offers a way of checking if these indicators are valid, uninformative or misleading.

This report concerns the practical implications of the current use of quality indicators by institutions to attract students, improve quality, improve 'value for money', and improve their relative standing in relation to educational provision. Institutions are responding to this data-driven market in a variety of ways, sometimes measurably successful, sometimes with negative consequences. It suggests ways in which the use of data in a market could be tuned up to have more positive effects. It identifies the changing nature of quality assurance and quality enhancement in this new environment and some organisational and quality side effects.

This report takes as its starting point the literature and evidence base of *Dimensions of quality* that identifies valid, uninterpretable and invalid teaching quality indicators. Where appropriate, the sources underlying *Dimensions of quality* are also cited here, but the sources and their interpretation will not be discussed again here.

It also draws on literature about the behaviour of public sectors such as health and education when they are obliged to act as if they were in a market driven by performance indicators (de Bruijn, 2002; Freeman, 2002; Propper and Wilson, 2003; Smith, 1990). Often, unintended and negative side effects are substantial, including what is often termed 'gaming' (efforts to improve indicators without attempting to improve quality) and vigorous attempts to improve invalid indicators that are used, at the expense of attempts to

improve performance in relation to valid indicators that are not used. Reviews of the literature tend to conclude that it is difficult to tell if performance improvements, if any, outweigh performance decrements. What is clear is that public sector organisations change their behaviour when performance indicators assume some importance, and early indications of changes in institutional behaviour in higher education are reported here.

The main sources of information for this report have been interviews and discussions with pro-vice-chancellors and others responsible for teaching and for quality assurance within 12 institutions of a wide variety of kinds, and examination of institutional documentation about change initiatives of a wide variety of kinds. An early draft of this report was discussed at the end of 2011 at a Higher Education Academy (HEA) PVC Network meeting and discussions were continued via email thereafter. Acknowledgement should be made of the openness with which these discussions took place. It does not yet seem to be the case that institutions are so competitive that they do not wish to share their ideas and practices with others. The institutions involved have not been identified.

2. Definitions of 'quality' underpinning student choice

While recent policy and funding changes may indeed influence the extent to which higher education is perceived by students as a 'market' in which they choose where to invest their fees, it is as yet unclear what features of the market will prove to be the main drivers of student choices.

Reputation has tended to dominate markets. In the US much effort is devoted to establishing a rank ordering of higher education institutions in relation to reputation, and academics are polled annually regarding their perceptions of reputation. However the resulting rankings are predictable on the basis of research performance, student entry standards and fee levels, together or separately, none of which predict student engagement or learning gains¹ (Graham and Thompson, 2001; Astin, 1985). Reputation measures are largely invalid as indicators of educational quality. Institutions with an existing high reputation may have a vested interest in resisting the introduction of more valid indicators of educational quality. In the US the Ivy League had initially been conspicuous in their avoidance of use of the National Survey of

¹ 'Learning gains' indicate the difference before and after studying. They are an important measure because a good deal of student performance is predicted by their prior educational performance, rather than by educational practices, while learning gains are predicted by educational practices.

Student Engagement (NSSE) that would allow comparisons with their less illustrious competitors (NSSE, 2007). Institutions with a lesser reputation but good educational provision may wish to emphasise valid indicators. Reputations are quite difficult to influence and may be unaffected by published quality indicators (as with the case of employers' out-of-date and ill-informed rankings of institutions). Harvard University's market position has not been noticeably damaged by its very public admissions that its teaching is not very good and Harvard students have been quoted as saying that they did not choose to go there for the teaching. If students are primarily interested in the behaviour of employers, then it is rational for students to ignore valid teaching quality indicators and focus on reputation.

Markets are driven to some extent by 'customer satisfaction'. The NSS and KIS emphasise student satisfaction. Ratings for various aspects of educational provision and facilities in the NSS are closely correlated with overall satisfaction, suggesting that there is a global halo effect around satisfaction that affects all other ratings. Satisfaction ratings for a programme can also change annually without changes in educational provision. There is currently no evidence concerning whether NSS overall satisfaction ratings are a valid indicator of educational quality. Institutions may be able to improve satisfaction ratings without needing to improve educational provision or quality. For example, the high Open University satisfaction ratings may be due, in part, to its student-facing and reliable administrative arrangements. It is clear that higher education institutions are trying very hard to improve student satisfaction, even when they are sceptical about its meaning or educational consequences.

Data about the proportion of students gaining 'good degrees' are difficult to interpret as standards vary between institutions (Yorke *et al.*, 2000). Also, current data about graduate employment and salaries are extremely difficult to interpret in trustworthy ways. This is due to the unreliability of the data and institutional variability in the way data are collected, to the short-term nature of current employability data, and to the raft of variables that affect employability that are unconnected with the quality of educational provision, student performance or learning gains. It is also unclear whether an individual student can expect to benefit in terms of increased employability and potential salary if they do not actively take advantage of an institution's, usually optional and extra-curricular, employability initiatives. Such benefits depend in large part on student behaviour rather than average employability data.

Several aspects of educational provision are known to predict both student performance and learning gains, independently of other variables such as resources, research performance and student entry standards. The most significant are class size, cohort size, extent of close contact with teachers, the

quality of the teachers, the extent and timing of feedback on assignments and the extent of collaborative learning (Gibbs, 2010) (see 'Implications for quality assurance' below for a fuller elaboration of these variables). The extent of access to appropriate learning resources also predicts student engagement and effort (Gibbs, 2010). The impact of these aspects of educational provision can be validly measured by examining student engagement and effort using the NSSE, and NSSE scores are sufficiently robust that they can act as proxies for the quality of educational provision (Pascarella *et al.*, 2010). With the exception of the NSS question about feedback, none of the above quality indicators are currently included in KIS, and neither is engagement or effort. In contrast, the proportion of class contact hours is included in KIS, but this is not a valid predictor of learning gains (Pascarella and Terenzini, 2005).

If the provision of new information to prospective students is to have positive leverage over educational quality then it will have to involve variables other than those currently available to students. Furthermore, if invalid indicators are used, institutions will change their behaviour to improve these indicators, sometimes at the expense of educational effectiveness. For example, increasing class contact hours, with fixed resources, often increases class sizes, reduces close contact with teachers and increases the use of part-time teachers, all of which are negative predictors of educational gains (Gibbs, 2010). When students make demands for increased class contact hours they may not be aware of the likely consequences.

3. Implications for quality assurance

Institutions have elaborate and expensive procedures for approving new courses and for reviewing existing courses, in order to assure quality. The adequacy of these procedures is audited by the QAA. What course documentation is required to contain in effect defines what aspects of quality are assured. If quality assurance pays attention to variables that do not influence quality and does not pay attention to variables that do influence quality then they are unlikely to assure or improve quality. The available research evidence on what variables affect student learning gains provides a fairly clear indication as to what such documentation should contain:

- **Class size:** planned cohort size for modules in the first, second and third years; class size and frequency of the smallest class students experience (such as a seminar) in each module, presented in a way that clearly indicates the extent of 'close contact' with teachers. It seems wrong that a student can commit fees for a degree programme without realising that they will be taught in lectures of 1,000 and in 'discussion' classes of 30.

- **Who does the teaching:** planned proportion of total teaching hours to be undertaken by full-time academics, by part-time or hourly paid teachers, and by students (such as postgraduates); proportion of all teachers contributing who are HEA-accredited teachers and/or who have a postgraduate qualification in teaching in HE. It seems wrong that a student can apply for a degree programme without realising that their discussion groups will be taught, and their work marked, largely by graduates who are not yet qualified as academics, let alone as teachers, or by hourly paid teachers who have no office and who cannot be contacted outside of class. Trained teachers are rated more highly by students, are more sophisticated in their thinking about teaching, and have students who take a more sophisticated approach to their studying.
- **Learning resources:** investment per annum per student, in the subject, in library and other learning resources; extent of availability to students of specialist resources (e.g. hours a week in the studio or lab).
- **Feedback on assignments:** how many times in three years students receive feedback on assignments; how many times this is 'formative only' with no marks contributing to degree classifications; total volume of feedback, in words; total volume of one-to-one, face-to-face feedback on assignments, in hours; average turnaround time for feedback, in days. Data on the current wide range of these indicators can be found in Gibbs and Dunbar-Goddet (2009).
- **Extent of collaborative learning:** proportion of total student learning hours in which it is intended that students study collaboratively, and alone, respectively.
- **Engagement:** actual total student learning hours, on average per week (the current range is about 18-36 hours) and in total for a three-year degree programme (the Bologna definition is 4,500-5,400 hours, the current UK range starts at below 1,500 hours (HEPI, 2007)); extent of students taking a 'deep approach' to learning (intending to make sense of material rather than only remember it) as measured by the Approaches to Studying Inventory (ASI: Entwistle and Tait, 1994), the Course Experience Questionnaire (CEQ: Ramsden, 1999) or the National Survey of Student Engagement (NSSE, 2007). Module documentation should indicate how the required volume of student effort is to be generated, rather than only focusing on what teachers do with their time. Most such documentation simply multiplies the number of course credits by ten and specifies the number of hours hypothetically involved. There is rarely any attempt to work out how these hours are generated by the required learning activity,

or even to specify what the learning activity consists of. KIS cites the proportion of class contact and study but the study hours are calculated by subtracting class contact from hypothetical total learning hours, and are not based on actual learning hours. There is rarely any attempt to evaluate how many hours students actually study, to check that course design assumptions are broadly accurate in relation to the demands that are made. Published evidence about students' actual workloads (HEPI, 2007) makes it clear that at present design assumptions about learning hours are overestimated by 100% in many cases: most UK students study for far fewer hours than KIS data suggest, dramatically changing the actual ratio of class time to study time.

It is common for institutional regulations about course documentation to require the inclusion of none of the above information.

Forms of course documentation require inclusion of certain important variables that oblige course designers to pay attention to them, but also omit important variables that allow them to be overlooked. This has allowed crucial aspects of educational provision, such as formative assessment of student work, to virtually disappear in many institutions. A recent study (Jessop *et al.*, 2012) of the impact of quality assurance on educational provision in the UK found that only one degree programme in 12 had any formative-only assessment at all in three years of study. Institutional quality assurance regimes have allowed this to happen and the QAA has expressed confidence in these quality assurance regimes. One institution in England has a ratio of formative-only to summative assessment of about 10:1, while in other institutions the ratio may be 1:10 or worse (Gibbs and Dunbar-Goddet, 2009; Jessop *et al.*, 2012). Variations in quality on this scale come about when quality assurance pays attention to the wrong variables.

It is not just quality assurance documents that fail to mention these key indicators of quality, prospectuses also usually fail to mention what kind of educational provision (for example, in terms of class size, close contact with teachers and feedback on assignments) students are likely to experience, and the KIS does not currently help. A market in which purchasers cannot find out what provision they will get for their money seems unlikely to provide much leverage to improve that provision.

In many institutions there is wide variation in NSS scores between subjects, to the extent that an institution may have the top ranked department in the country in one subject but the bottom ranked department in another subject. This suggests that current institutional quality assurance regimes do not, at the moment, have a very influential role in levelling up quality

across departments and subjects. In a few institutions there are poor NSS ratings across all programmes, and no strong subjects, suggesting that the quality assurance process is either focusing on the wrong variables and using the wrong levers, or may even operate to limit the scope subjects have to improve. Bureaucratic change processes (see pp.32-34 and Case study 3) may exacerbate this problem.

4. Implications for cost-effectiveness

There has been a relentless decline in funding per student over an extended period and it is as yet unclear whether new funding regimes will ameliorate or exacerbate this situation over time, though it seems likely that it will ameliorate it a little for some and exacerbate it markedly for others. What seems inevitable is that the pressure to operate cost-effectively, and at the same time to offer 'value for money' for students, will continue. Competition from new providers of HE who do not carry some of the costs associated with conventional HE institutions will sharpen this emphasis on cost-effectiveness, especially if they are able to offer more extensive educational provision for a lower price.

Using funding to increase class contact is, on its own, unlikely to improve education, whatever the students' unions demand. Total teaching hours do not predict student performance and increasing total teaching hours may have the consequences of increasing class sizes or simply reducing study hours, both of which would have negative effects on student performance. More important is the volume of student effort generated by whatever pattern of teaching and related assessment is provided (and supported by whatever learning resources are available). Cost-effective pedagogic regimes generate much productive studying with modest teaching input (and can still receive very high student feedback ratings, as at The Open University and the University of Oxford).

Students are a much underutilised resource in that pedagogic processes that give students an active role, such as collaborative learning, peer tutoring and self- and peer assessment (for feedback, not marks) all increase performance at little or no cost. Improving students' effectiveness as learners has more impact on performance and learning gains than does improving teaching or improving curricula, and a great deal more impact than organisational change (Hattie, 2009). The role of feedback in learning is a good example of this phenomenon. Institutions with poor NSS scores for feedback often take steps to try and increase the volume of feedback that students receive. This may be valuable but it is expensive, and there is no simple relationship between the

volume of feedback students receive and the extent to which they read it and find it useful. In contrast, successful students use feedback differently, more extensively and more effectively than do unsuccessful students. Changing students' use of feedback can be much more effective than changing the feedback, and may involve no costs. It is also the case that increasing students' social engagement improves retention and increases their self-efficacy (their sense of being able to succeed) and improves their performance. Developing the sophistication and effectiveness of students as learners is a worthwhile goal of higher education, but it is also justified on cost-effectiveness grounds.

Summative assessment for marks is hugely expensive in both academic and administrative time, costing more than teaching in some contexts. It has increased as a consequence of small sized modules, each of which is required to generate a mark, the proliferation of module-level learning outcomes, each of which is supposed to be assessed and marked, and the shift from examinations to coursework assessment. The number of times student work, which contributes to their degree classification, is marked is ten times greater in some institutions than in others (Gibbs and Dunbar-Goddet, 2009). Frequent formative-only assessment, with quick feedback and no marks, can be cheaper and also much more pedagogically effective. Greatly reduced summative assessment could usefully be concentrated on programme-level outcomes, towards the end of programmes. Larger, longer and linked modules can also reduce the need for current volumes of summative assessment. Making formative-only assignments compulsory can guarantee that students tackle them, though current quality assurance and examination regulations sometimes forbid this practice. The pressure now is more likely to come from students who may want all their assignments marked, in return for their fees, whether or not this will improve their learning.

If student effort is to be generated cheaply, it has to involve good access to learning resources. As The Open University demonstrates, investing heavily in learning resources reduces the need for teaching and reduces overall costs without compromising quality. The Open University charges the lowest fees despite being top of the NSS rankings. Library investment per student predicts student study hours (Gibbs, 2010). Teachers' generating their own resources for digital storage is a very expensive way to provide access to learning resources. Teachers could instead be designing learning activities that generate productive learning effort, wrapped round existing accessible learning resources. In general there may need to be a lesser emphasis on teachers providing content (in class or digitally) and a greater emphasis on teachers structuring the learning processes through which students engage with the content, and teachers then giving feedback to students on their learning. In the US there is currently much interest in developing what they

term 'flipped' courses to increase cost-effectiveness, in which class time is used to brief and debrief extended study out of class supported by learning materials, rather than using scarce class time to present material in the hope that students might then study it out of class. Increasing class contact (and also increasing the number of rooms full of PCs) reduces the availability of flexible learning spaces for students to engage with the content.

There is ample evidence from the US that student engagement and learning gains can be increased despite reductions in resources (Twigg, 2003), but little sign that this expertise is being drawn on in the UK. It is now more than two decades since the 'Teaching More Students' national initiative trained nearly 10,000 teachers how to cope with large classes with educationally sound methods (Gibbs, 1995). There is a need for a new national initiative on cost-effective teaching so that, where fees are lower, or where income for teaching has declined for other reasons, forced changes to teaching practices could make things better rather than worse.

5. Implications for quality enhancement

While many institutions are embarking on quality enhancement initiatives with an intensity not previously seen, their efforts are not always targeted on variables that are known to improve student learning gains, but only on improving those quality indicators that are in the public domain.

Training teachers improves students' ratings of teachers and improves teachers' sophistication (Nasr *et al.*, 1996; Gibbs and Coffey, 2004). If part-time and hourly paid teachers are used to a significant extent then they need to be briefed and trained for the specific roles they play, and their effectiveness monitored. Increasing the use of untrained part-time teachers compromises quality and reduces student performance and retention (Ehrenberg, 2006). It is not known if the proportion of teachers who together teach a programme, who are trained, influences learning gains or student ratings for that programme. The HEA has data on accredited teachers and NSS scores and could test this relationship.

Departmental leadership of teaching makes a difference and should engage teachers in talking about teaching, creating a functioning 'community of practice' that shares values and approaches. Creating degree programmes from modules and teachers from disparate departments makes this much harder to achieve. Matrix management structures also limit the scope for functioning communities of practice. However, some institutions are paying more attention to the training and support of programme leaders, and even preparing future programme leaders.

It helps to collect valid evidence before and after attempts to improve – e.g. using scales from the NSSE about student engagement, and about levels of student effort – rather than relying on student satisfaction or ‘home-made’ feedback questionnaires, which are both unreliable and poor predictors of educational quality. If institutions are to be required to use and publish student feedback then it makes sense to use valid and reliable questionnaires that make comparisons interpretable, rather than inventing one’s own.

It helps to focus more on improving *programmes* and their coherence and progression, consistency of demands on students, and on the assessment of well-understood, programme-level learning outcomes. Focusing on the (diverse) improvement of individual modules and individual teachers can reduce coherence and make programmes work less well, even when individual teachers and individual modules are quite good. It is programmes that students rate on the NSS and poor overall ratings can be generated despite good student experiences of some individual modules and teachers. For example, a national R&D programme concerned with improving assessment and feedback (TESTA) has found that diversity of assessment between modules within a programme was a frequent cause of students not using, or valuing, even extensive feedback – because the next assignment they were asked to do was too different for the feedback to be able to feed forwards. Solving such problems requires collaboration between module designers across modules and years and coherent programme design.

Career structures and financial rewards for academics are often heavily weighted towards research and administrative achievement, and may tolerate the neglect of teaching. Institutions that have improved overall in their teaching in relation to other institutions in their country (such as at the University of Utrecht and University of Sydney) have radically changed their career structures and promotion criteria, making the neglect of teaching a very risky personal strategy and making a career in ‘leadership of teaching’ attractive for able academics. Educationally successful institutions value teaching in overt ways and have developed ways to judge teaching with a similar level of confidence to that used in judging research. The priority should not necessarily be to reward the best teachers, but to make the baseline level of expectation much higher. Stanford University, and MIT, for example, very rarely give tenure for outstanding teaching, but instead are very careful to avoid giving tenure to outstanding researchers who are not also, at the minimum, good and conscientious teachers.

Awards and public acknowledgement of special achievement should be reoriented towards recognition of outstanding programmes and ‘learning environments’ that require the collaboration of many teachers, and away from public competition between individual teachers.

6. Reviewing teaching quality indicators annually

It has become common, perhaps ubiquitous, for there to be some kind of institutional post-mortem after NSS results are published each year. In many cases this is not left up to subjects or departments. The centre will collate NSS data and add other institutional data such as about entrants' A-level scores, applicants per place, retention rates and marks in examinations, and send a report to each department for consideration. In some cases this can be quite an elaborate procedure. At one university, for example, each department identifies ten benchmark or competitor institutions that teach their subject, and a new office at the centre collates a special report for each department allowing them to make comparisons with these ten competitors. To NSS data are added data from other surveys, including international surveys, HESA, HEFCE, HEPI, etc. The extent to which such reviews are data driven, and involve careful scrutiny of quality indicators from competitors in the market, is quite different from any previous review process. PVCs have described previous review processes as 'only qualitative', or even 'fluffy', in comparison with the harder data-driven approaches now used. Departments are described by PVCs as far more aware of who their competitors are and exactly how they stand in relation to them.

Some institutions have set up special groups with responsibility for identifying potentially useful sources of quantitative data linked to institutional goals, or initiating the collection of new data, for example data from employers of graduates to inform missions concerned with employability.

Such data-driven reviews often now take place annually rather than only periodically, perhaps as seldom as once every six years, which used to be the case. In some institutions this new annual review has in effect supplanted periodic review except for the formal role review performs in relation to the QAA and its quality assurance expectations.

In some cases, where the quality indicators are especially poor, or where they have been poor several years running, subjects are being closed down: something that rarely happened as a result of periodic review. Whatever the actual reliability or validity of NSS and other data, they are taken more seriously and given more weight than the kinds of observations previously made about the quality of programmes. It seems as though the content and focus of the NSS has trumped the agenda and concerns of the QAA without much debate having taken place.

It is not only that this new focus of reviews involves data to an unprecedented extent. It also often involves meetings, both among teachers within a department and with the PVC or head of quality. Characteristically there is a face-to-face meeting after the department has had time to digest the data. In several institutions departments are obliged to hold a departmental meeting, of a day in duration, at the end of which the PVC turns up to discuss their interpretation of the data. The PVC may be accompanied by the head of the teaching development unit and the department may be obliged to come up with some kind of developmental response: how they are going to address issues that the data exposes. An oral statement at the meeting might then be required to be formalised in a written plan, which is used as the basis of a similar review and discussion the following year.

In some cases these meetings and discussions appear not to be collegial. A team of senior academics may descend on a department with poor quality indicators, whether the department wants the meeting or not, and whether they value the quality indicators or not. This kind of selective focus on 'failing' departments (with a 'light touch' for departments with acceptable quality indicators) echoes changing approaches to quality assurance for schools and the QAA's proposed risk-based approach to institutional audit. It is quantitative data that are determining who gets the light touch and who the heavy hand, and the content of that dataset has often not been negotiated within the university. Heads of department are often highly sceptical about the value of the data that are being 'used against them'.

Analysis of NSS data also extends to service providers (such as computer services) who now have benchmark data from competing institutions against which to judge institutional performance. Interpreting such NSS scores is proving difficult and often requires greater exploration, both quantitatively through further data collection or targeted surveys, or qualitatively, through the active involvement of students in discussing what the data mean, before a credible diagnosis of causes of problems is achieved. As with quality indicators about teaching, the data are often perceived as too coarse-grained to provide much insight, but succeed very well in prompting more evaluation and attention than might otherwise have taken place. Efforts to improve services (or at least students' perceptions of services) seems more likely to improve quality indicators generated locally that concern local phenomena, than to improve NSS scores, which are often perceived as too general to reflect detailed improvements, or

even as somewhat intractable. When asked about whether efforts to improve NSS scores for services had succeeded in improving these scores, PVCs often expressed the view that this was an unrealistic goal.

Almost all institutions had student feedback questionnaires, course review meetings, student representatives on course committees, and so on, long before the NSS was implemented. What is new has evolved from these standard features of quality assurance. Case study I, set in a new university, outlines such an evolution in three distinct phases.

Case study 1: An evolving student feedback system

In Phase 1 all modules used a standard module feedback questionnaire. It was administered at the end of the module and so could not influence the current operation of the module. It involved questions and a focus of attention that differed from the NSS, was not accompanied by any required review of the data, and was widely ignored.

In Phase 2 the module questionnaire was redesigned to mirror the NSS. It was administered by students who visited classrooms during Semester 2, and the data were collated and processed centrally. There was a ten-day turnaround time to get the data back to module leaders, who were given comparative data. Compulsory observation of all teachers by a more senior colleague was introduced, with pre- and post-observation meetings, usually developmental and private. A five-day time limit was set for module leaders to produce action plans in response to the data. Each department drew up a 'league table' of questionnaire scores for modules and the centre drew up a 'league table' of departmental scores. Poor scores were therefore made very public. A 'traffic light' system was devised to identify poor, or very poor scores, with the lower boundaries gradually moved upwards over time. Typically each Dean would then meet each module leader individually to discuss their scores and action plans. The institution's NSS 'overall satisfaction' score increased from 79% to 85% in the first year of operation of this scheme. Policy changes were made centrally that affected all modules, such as requiring all assignment feedback to be reduced from five weeks to three weeks, with a future target of two weeks. NSS feedback scores improved by 15% and the institution moved from the bottom quartile of scores concerning 'feedback' to the top quartile, in the first year of implementation of this policy.

In Phase 3 the focus shifted from modules to programmes, both for the collection and consideration of student feedback data and for course approval and review. Each team of teachers delivering a programme is required to hold two one-day meetings a year. At one of these meetings they are required to discuss:

- a set of data generated centrally, including NSS data;
- the 'dominant pedagogy' of the programme;

- the overall assessment strategy of the programme, including the proportion of assessment devoted to programme-level learning outcomes (as contrasted with the previous emphasis on module-level outcomes);
- the institution's learning and teaching strategy in relation to the data, for example the extent of take-up of opportunities for international work experience.

Part of this meeting is attended by senior management, normally the PVC. After the meeting the programme team is required to produce a report (which replaces the previous monitoring report, which staff perceived as only a bureaucratic requirement) setting targets for key quality indicators for future years, and an action plan for reaching these targets.

This focus on programmes has been accompanied by a reduction of 50% in the number of programmes offered, so that quality assurance and enhancement attention and effort can be more focused, and so that programmes become more coherent. It is too early to measure the impact of Phase 3 on quality indicators.

7. Relating programme-level quality indicators to collections of courses

Since the very wide adoption of modular course structures, starting in the 1980s, quality assurance and enhancement has tended to focus on the module, rather than on the programme. Quality indicators, that potential students may see and pay attention to, focus on programmes, not on modules or the quality of individual teachers. To improve NSS scores you have to focus your attention on programmes. This is posing a number of challenges.

There is a weak relationship, in some contexts, between the scores that are published for a 'subject' and the collection of modules that students actually experience. For understandable practical reasons published data are collated under conventional subject labels, but these often do not correspond with local definitions of a subject. In their response to NSS questions, modular course students are being obliged to construct a single rating to encompass their experience of disparate modules. A number of loosely related clusters of modules may be lumped together and averaged. Many degree

programmes at an institution may be represented in published rankings by a single NSS score for a subject that is not actually taught. Students on modular courses also choose to study modules from disparate clusters and their personal experience cannot in practice be easily unpacked into its component 'subject' parts. It is difficult for modular institutions to make sense of the data, and programme teams may have trouble making sense out of ratings to which their collection of modules may have only contributed 30% of the student experience. There is also less incentive to improve as programme teams may have little control over the other 70% of data that contributes to published scores, and even improving their 30% might not make much difference overall.

A related difficulty is that some of the large-scale developments in institutional responses to the current market involve activities outside of degree programmes (such as efforts to engage students through various entrepreneurial activities, internship schemes, or many institution's generic efforts to improve employability). In the US, successful efforts to improve student engagement, as measured by the National Survey of Student Engagement, are commonly extra-curricular. In some UK institutions with traditional single subjects, the link between such engagement-oriented efforts and the degree programme is quite close: they take place inside the department and involve the same academic staff. In modular institutions such efforts are often entirely separate from the clusters of modules that students rate on the NSS. It is unclear the extent to which students' appreciation of such extra-curricular activities is reflected in NSS scores that refer to a specific programme or clusters of modules.

One consequence is that institutions with conventional degree programmes that correspond closely to NSS subject definitions are at a clear organisational advantage. Other institutions are starting to reconfigure modular programmes in ways that make NSS scores more interpretable and engage those who teach on the programme to a greater extent in examining and using the data.

8. Retreating from modularity

There are strong indications that there is currently a rapid retreat from modularity. Clusters of modules are being put back together into more coherent subjects with more conventional labels. The scope for students to take modules from unrelated areas is being curtailed. A number of institutions have cut the number of allowable combinations of subjects by 50% within a single year. Modules are being made larger (in relation to credits or study

hours, or spanning semesters), reducing unhelpful variation between modules in their pedagogy and teaching and assessment practices and offering scope to reduce summative assessment. Assessment regimes are being overhauled so that there are vertical links, between years, allowing students to perceive coherence from year to year and to progress in the sophistication with which they use the discourse and skills of the discipline. There is a much greater emphasis on assessment of programme-level learning outcomes, in some cases replacing module-level outcomes altogether. These programme-level outcomes include some of the 'USPs' institutions are emphasising in their marketing, such as the development of skills valued by employers.

There are a number of reasons for this retreat, but the focus of the market on programmes appears to be accelerating the change. PVCs have also expressed a need to reduce administrative and other costs associated with complex course provision (and accompanying examination arrangements) and focus these resources and effort on 'core' programmes so that there is more likelihood of measurable improvements in quality.

Many of these changes are led, or mandated, from the centre, with departments working out how to implement the organisational changes locally, rather than the changes growing out of curriculum reviews within disciplines. For example, at one university the modular system has been abandoned entirely and in all subject areas students now study just three courses: a first-year course, a second-year course, and a third-year course. All teachers in all subjects have had to work out how to put their modular degree programmes back together again within this imposed framework. This may or may not improve quality, but it is not how curriculum development used to function.

The overall effects appear to be:

- what constitutes the student experience of a degree programme is much easier to control and understand, and NSS scores are much easier to interpret and act on;
- coherence and progression may well improve, and with it the quality of student experience;
- teachers are talking to each other more about the way their module fits in with, and contributes to, the whole. There is more discussion of shared goals and of programme-level learning outcomes. There is more team teaching on the larger course units that have been created, and fewer 'silos' where isolated individual teachers pursue personal interests;

- student choice is being curtailed, with the inevitable likelihood of students studying subjects they are less interested in and having less scope to pursue subjects they are interested in. The negative effects of this on students' ratings may be difficult to unpack from their positive ratings due to greater coherence. Institutions may need to add evaluation questions of their own to the NSS to explore such issues;
- curriculum innovation, in relation to cross-, inter- and multi-disciplinary arrangements, is being turned back towards traditional single subject disciplinary definitions and boundaries. If there is a value of learning at disciplinary boundaries, or in new subject configurations, this is sometimes being sacrificed due to the way NSS data are currently collated and reported. The ability to compare courses using standard quality indicators, across institutions, is being given a higher priority;
- where reducing the number of course options increases the number of students taking each option, educational effectiveness is likely to be reduced: by increased cohort size, by increased class size and by the increased use of part-time teachers necessary to teach the larger number of discussion classes and mark the increased number of assignments and examination papers.

9. Focusing quality enhancement on programmes and course teams

A review of the way teaching development efforts in universities, worldwide, have evolved over recent decades (Gibbs, 2013) has highlighted a shift in attention from a focus on individual teachers and individual course units to a focus on course teams and degree programmes. Case study 2, set in a new university, illustrates difficulties associated with not making this shift in a market in which quality indicators are about programmes.

Case study 2: Focusing on teachers and modules

This University had an enviable record of success in winning National Teaching Fellowships and a spectacular record of earning external grants to develop teaching, such as from the Centres for Excellence in Teaching and Learning initiative. It matched this external success and visibility with internal processes: extensive training of individual teachers, starting decades ago, internal funding for innovation in teaching, teaching awards for outstanding teachers, seminars for enthusiastic and innovating teachers, an annual event to showcase innovation in teaching, and so on. In relation to the kinds of quality enhancement mechanisms that the QAA like to see when auditing a university, it looked very strong.

However, its NSS scores as an institution have been moderate and it has many 'below average' subjects. The quality peaks are not sufficient to balance out the quality troughs.

The explanation appears to lie in the University's strategy of reliance for improvement on individuals, who teach innovatory modules, that succeeds in engaging only a relatively small proportion of enthusiastic teachers but which leaves much of the University untouched. It also lacks any consistent direction: innovations may make students' experience inconsistent across a programme, or may be oriented to goals other than departmental or institutional goals. Even expensively funded initiatives in this institution have sometimes faded away once funding stopped, failed to spread beyond the department or even failed to engage the departments in which they were based.

The University has no teaching awards for outstanding degree programmes, little in the way of leadership development programmes for programme leaders, no rewards for those who succeed as programme leaders, and no career structure that would encourage anyone to take on 'leadership of teaching' duties (indeed promotions are largely based on research achievements). It has few mechanisms to spread successes sideways. The emphasis has tended to be on trying something new rather than widespread adoption of existing successful practices. Quality assurance has involved a good deal of documented detail about individual modules but less about programmes. There are few mechanisms to address marked weaknesses in programmes, some programmes have been perceived as moribund and unengaged for a

long time, and a number of subject areas have little history of engaging in the institution's teaching development mechanisms.

As other institutions, including their nearest competitor, climb the teaching rankings through institution-wide strategies and a clear focus on departments, this institution is being left behind.

The shift of focus of quality enhancement to the unit of analysis of quality indicators, the programme, is not yet strongly in evidence, though there are signs of the shift in a variety of other changes.

As discussed above, NSS scores and other quality indicators are being discussed by course teams, often compulsorily, in ways that did not previously take place.

As discussed below, enhancement effort is being redirected away from idiosyncratic efforts by keen teachers towards comprehensive implementation of programme-wide changes, driven by programme quality indicators. Progress is being judged by programme-level quality indicators, rather than by listing individual teacher achievements or describing isolated innovations.

The higher levels of the HEA's accreditation scheme (Senior Fellow and Principal Fellow) are receiving attention and there are a number of institutions developing training, support and professional development frameworks for those who lead change in teaching, either through supporting the development of a team of teachers or through curriculum development. This attention is not restricted to the early adopters of the Staff and Educational Development Association (SEDA) and Institute for Learning and Teaching (ILT) teacher accreditation schemes, but includes research-led universities. There is not yet widespread evidence of serious development of leadership of teaching of the form found at, for example, the University of Utrecht, or as recognised in career structures, as for example at the University of Sydney.

Some institutions are developing roles, and associated job titles and salaries, for those that lead and develop degree programmes.

Over the past decade there has been a considerable emphasis on developing mechanisms that reward and recognise excellent teachers, through awards,

through promotion, through recognition schemes and special job titles, through special funding for innovation, and so on. National initiatives such as the National Teaching Fellowship Scheme and the strand of Centres for Excellence in Teaching and Learning (CETLs) concerned with reward have also focused almost exclusively on individuals. However, NSS scores and other teaching performance indicators are not about individuals. There has been no corresponding emphasis on rewarding excellent programmes or departments that create the learning environments to which students respond. In contrast there is a national scheme in both Norway and Finland that recognises outstanding teaching departments, and this has spawned corresponding institutional award schemes such as at the Universities of Oslo and Helsinki, that focus on teachers collaborating in curriculum design and pedagogy so as to make learning environments especially effective. Similarly the University of Utrecht has a single, highly prestigious, teaching award, presented alongside honorary doctorates each year, for the outstanding 'leader of teaching' who has galvanised others to change whole programmes collaboratively.

10. Adopting new institutional change processes

Quality problems identified through NSS scores or other quality indicators have led a number of institutions to make institution-wide changes, short-circuiting any 'emergent' change process of gradual innovation in teaching and sideways spread of good practices, and cutting across departmental autonomy. Some of these institutional changes appear larger in scale, or even draconian, compared with past approaches to change in teaching and assessment. Where radical changes have been implemented comprehensively, and then accepted, this provides an indication of the changed environment that teachers now perceive themselves to be in.

Two broad categories of change process appear to be common, characterised well by McNay's (1995) nomenclature of organisational cultures in higher education: 'bureaucratic' and 'corporate'². The two cases outlined here illustrate the nature and scale of change that is evident. Case study 3 is based in a large arts institution.

- 2 McNay distinguishes four organisational cultures in higher education: collegial, bureaucratic, corporate and entrepreneurial. Collegial cultures rely on individual and departmental autonomy, for the quality of teaching. Bureaucratic cultures attempt to bring about change through imposing policies and regulations. Corporate cultures attempt to bring about change by orienting the whole institution to corporate goals. Entrepreneurial cultures emphasise professional expertise and dispersed leadership, accountable through market position.

Case study 3: Centrally driven change to assessment criteria across all programmes

It is common in arts institutions for assessment criteria to be tacit and weakly articulated, and to differ between sub-disciplines in ways that only 'insiders' can understand. NSS results at the institution that is the focus of this case study had consistently showed that students did not perceive assessment to be fair and did not understand criteria, and there were also relatively poor scores for 'feedback'. With up to 50% of teaching undertaken by hourly paid staff, and a lack of social spaces for staff, it seemed unlikely that shared standards based on collegial processes involving informal interaction among staff could in practice be established satisfactorily. There was also evidence suggesting that students from under-represented backgrounds found it harder to deduce tacit standards, with adverse impact on their achievement, and evidence of a gap between the achievement of white students and BME students wider than the national average.

The institution had initially adopted a collegial approach to this problem, encouraging teachers to address it in their own ways, but without rapid progress. It then addressed this situation bureaucratically by imposing a standard set of eight explicit marking criteria, a standard feedback form that has been developed into an online form integrated with the student record system, and a matrix indicating characteristics of achievement corresponding to each grade. There are undergraduate and taught postgraduate versions of each set of criteria but no variations between assignments, subjects, degree programmes, or years of study. Courses could adapt the feedback form by indicating any criteria that are not applicable, and by customising the explanatory text appearing below each criterion to make it more relevant to their students, but they could not ignore it. The matrix (in effect determining programme-level learning outcomes) could not be adapted: it had to appear in its standard form in all course handbooks.

The purpose of the initiative was to introduce explicit standards into assessment, as the basis for marker moderation discussions and for explaining assessment to students, to prompt more effective feedback on student work and to speed up the process of providing such feedback for markers, and to underpin course design. Feedback from staff suggests that the development has at least partly met all these goals.

As this change was driven by quality indicators it was important that quality indicators were seen to improve as a consequence. In the first year after the change there was an average increase of 7% in the institution's NSS scores for assessment and feedback, and an increase of 7-10% for the assessment and feedback questions in the Postgraduate Taught Experience Survey. Student feedback obtained through focus groups was positive from the outset about the criteria.

Changes of this kind, that override local cultures and practices concerned with disciplinary goals and standards, have in the past been rare in arts and humanities contexts and even more rare across specialist arts institutions. It seems likely that institutions that see such changes working, at least in relation to NSS scores, are more likely to also impose such changes in the future. It may be the case that only institutions with poor quality indicators have the scope to improve through such change mechanisms, and that institutions that are already reasonably good need much more sophisticated and contextually nuanced change mechanisms to improve further.

Not all such centrally driven changes work, of course. For example, a number of institutions have, in response to poor NSS ratings for feedback, established an institution-wide policy of all feedback being returned to students within, say, three weeks. Quite apart from the fact that three weeks is actually rather slow if you want feedback to be useful to students, there is little evidence of institutions having any way to police such policies, or even evaluating their impact on assignment turnaround times, and students' experience, and their NSS ratings, may remain largely unchanged. Improvements in students' perceptions as to whether feedback is received 'in time to be useful' seems more likely to be achieved through changes to the timing of successive assignments, so that feedback from one can be useful in tackling the next and 'feed forwards'. Such local changes are too varied in nature to be mandated centrally.

A number of institutions have imposed a teaching method across the board in order to address a problem identified in NSS scores, or to strengthen an already successful feature of institutional pedagogy. For example, one institution has introduced a requirement for all students in all subjects to have a weekly small group tutorial whose function spans all the courses the students in the group are currently taking. This may well, among other things, increase 'close contact' with teachers, and research evidence suggests that is likely to increase student learning gains and retention (Pascarella and Terenzini, 2005)

(and this institution does indeed greatly outperform its benchmark for retention). However, this kind of direct intervention from the centre to impose a particular way of going about teaching has not until recently been a common phenomenon.

Case study 4 illustrates a centrally driven but locally implemented change across a whole institution, of a kind that would probably not have taken place a decade ago, in the context of a Russell Group university.

Case study 4: Student-driven change across all departments

The institution in this case study is research-led and in the past had a largely collegial approach to the development of teaching: responsibility had been devolved to departments and central initiatives had been limited in scope and had a limited history of successful implementation. Change initiatives now start with identification of an issue through the careful analysis of teaching and learning quality indicators, and involve a combination of central funding and co-ordination with departmental implementation. The initiatives involve extensive support from graduates who are from the departments themselves, and so are known locally and understand the local disciplinary pedagogy, but are employed, briefed and trained centrally. For example, a moribund virtual learning environment (VLE) was replaced by a new system, but with students extensively trained to provide technical and pedagogic support to individual academics to enable them to migrate existing course materials and learning materials to the VLE, and to exploit its many facilities such as online assignment submission and feedback. Within a year every course and every lecturer in the institution was using the new VLE and exploiting far more of its capacities, for example a 100% increase in digital recording of lectures and a 300% increase in other learning resources stored on the VLE. Both teachers and students were extremely positive about the educational benefits involved.

A series of large-scale university-wide initiatives have now adopted the same approach to change: central funding, central co-ordination and support, with trained student facilitators, targeted on issues identified through NSS scores or other quality indicators, but implemented locally in discipline-relevant ways, with centrally co-ordinated evaluation of impact on quality indicators. The impact on quality indicators at this institution has been very marked: the University has climbed the NSS rankings every year and currently has the best teaching quality indicators of universities of its type. This appears to be a more effective change strategy than that described in Case study 2.

II. Focusing on 'hygiene' factors and consistency of delivery of services

Some institutions have markedly higher NSS ratings for 'organisation and management' than do others. Institutions seem to vary in the consistency and adequacy of their administrative arrangements for things such as timetabling, room bookings, return of assignments and reporting of marks. Some institutions place a very high emphasis on student-facing administrative efficiency and collect a good deal of data on, for example, speed of response to telephone enquiries, or what proportion of tutors initiate online communication with their students before a course starts, and act on these data to improve the reliability and standard of all aspects of services. Some of this has immediate educational and financial pay-offs. For example, at one university it is known that if a student misses an examination but is offered *within 24 hours* a date for a resit then this greatly increases the chance of the student taking a resit, succeeding in the examination, and not dropping out of the university, with its associated loss of future fee income and increase in recruitment costs to replace the student. This university undertakes analyses of the costs of interventions and their financial benefits, to help decide which services to fund and mandate across the whole university to specified standards, rather than leaving it up to varied local arrangements (Gibbs *et al.*, 2006). Such an approach to the management of service delivery seems likely to become more common, and PVCs who were interviewed for this report often gave accounts of centrally driven initiatives to address perceived weaknesses in service delivery, driven by data about reliability or at least by data about students' satisfaction with services.

'Organisation' also emerges from factor analyses of generic student feedback questionnaires (such as the highly reliable 'Student Evaluation of Educational Quality' questionnaire: Marsh, 1982; Coffey and Gibbs, 2000) as a close correlate with 'good teaching' overall, and as a predictor of student performance. Good teaching is perceived by students to include such things as provision of clear briefs for assignments, reliable information about access to learning resources (the books are actually in the library and the websites are still live), handouts delivered in time to be useful, and so on. Getting these 'organisation of teaching' factors right can improve students' overall perceptions of the quality of teaching and in this case it is not just 'customer satisfaction', it actually improves learning.

PVCs are much more likely than in the past to take action when an individual department shows up unexpectedly badly in the NSS. This may involve a rapid and in-depth review of what had gone wrong, sometimes associated with additional collection of evaluation data, interviews with students and internal

departmental discussions, sometimes involving students. It has become clear to many PVCs that when large year-to-year variations take place it is often the case that the underlying curriculum, teaching, teaching staff and assessment had remained largely unchanged, but students' reaction to it had changed markedly for the worse, across most or all scales on the NSS: for example, ratings for feedback may decline despite no changes in feedback arrangements and practices. The underlying cause is often to be found in a gross administrative malfunction: for example, a collapse of work placement arrangements two weeks before students had to submit a work portfolio for assessment and one week before the NSS was administered. Students may suddenly become very annoyed and deeply unimpressed by the competence of the department, and express their annoyance by rating everything down. Often the unexpectedly poor ratings recover the following year with little or no change in any aspect of teaching or assessment, once the administrative malfunction has been addressed. The sometimes dramatically unstable rankings of institutions within individual subjects in relation to their NSS scores suggest that this is a common phenomenon. All kinds of things can go wrong that upset students' overall judgements, and the following year they do not go wrong and ratings recover, leading to yo-yoing of scores and rankings.

This is often an issue concerned with avoiding reputational damage rather than achieving teaching excellence. Where things can be seen to be going wrong across a number of courses, or regularly, it has become much more common for institutions to embark on quite wide-ranging initiatives to improve the overall reliability of delivery of all kinds of services across all aspects of institutional operations. In addition departments are becoming much more aware of the negative consequences of getting something administrative horribly wrong, and either nipping it in the bud or taking firm action to avoid a repeat. Idiosyncratic teachers who ignore policies on principle and who are poorly organised may be tolerated less.

12. Changing students' role from 'consumer' to active participants

Much of the rhetoric about the value of a higher education market treats students as purchasers, customers or consumers. These terms are anathema to much of higher education, not because higher education is anti-capitalist but because teaching and learning are seen as symbiotic. Students do not consume knowledge but construct it in a personal way in the context of learning environments that include teaching: they are co-producers and collaborators. The need for students to be active participants in teaching and learning processes is part of the *Seven Principles of Good Practice in Higher Education*

(Chickering and Gamson, 1987), and student engagement (as active learners) is the central focus of much quality enhancement in the US at the moment, measured by the NSSE.

However, there is a second and very different focus to student engagement that is currently gaining momentum, which is not about pedagogy but about change processes. It involves students as evaluators and active participants in educational decision-making as institutions improve quality. It is perhaps best embodied in initiatives that employ students as change agents (cf. Dunne and Zanstra, 2011) and build them into quality improvement initiatives that have changed educational practices. Some of these developments are driven by a belief that students will respond more positively to their experience (and rate it more highly) if they are actively involved in evaluation rather than simply questionnaire respondents. A number of institutions have in effect handed over the administration of local questionnaires and interpretation of the NSS to students, and there is some evidence that NSS scores improve when students are engaged in this way. Other developments involve a more sophisticated understanding of educational change processes and their greater likelihood of success and impact if students are actively engaged with them.

13. Competing research agendas

The effect of the Research Excellence Framework (REF) and its predecessors on institutional priorities and academic practices is well documented and understood. It affects values and attitudes towards teaching, it orients career and salary decisions, it changes how academic staff use their time and increases the use of research students to undertake the 'close contact' that is necessary for a good quality of student experience. This has had different kinds of impacts on educational quality in different contexts.

In research-oriented universities below the top rank, research selectivity has inevitably resulted in research income declining, while the level of research effort required to earn research income has increased. For some a point comes where the main competition, in relation to attracting good students prepared to pay top fees, is not the research elite but more teaching-oriented universities whose NSS scores may be as good even if their research performance is not. It is not that uncommon for very recently established universities to be the first choice of the majority of potential students from the local region, even when there is a prestigious research-oriented university next door. As the previous effects of research-based reputation decline, prospective students make more varied choices and the comparative

quality of student intake may decline, requiring greater teaching effort. In such institutions the risks of continuing to pursue research goals *at all costs* are becoming greater and there is a rebalancing of priorities going on. The kinds of centrally driven (and sometimes successful) teaching development initiatives now evident at universities ranked for research just outside the top dozen or so would have been unlikely to have been attempted a decade ago.

In some institutions, mainly polytechnics that became universities, the attempt to compete in research selectivity exercises moved priorities away from teaching, and in some of these institutions the way new academic staff are appointed, and the way promotion decisions are made, still emphasises research even at a time when the necessity to compete on teaching quality indicators is very pressing.

Some institutions have recently had to emphasise their research in order to achieve university status and doctorate awarding powers, even though they are clearly teaching-focused institutions, and they have continued with this research emphasis after gaining university status. In some such institutions it is difficult to implement comprehensive teaching development strategies, or even teaching development projects, because there are few rewards for good teachers, only new demands. In particular the need for leadership of teaching in roles such as 'programme leader' for a degree programme is not matched by the benefits to teachers of taking on such roles or committing to a career involving 'leadership of teaching'. A number of teaching universities have reward and career structures that must have powerfully damaging effects on efforts to develop teaching and improve teaching quality indicators. Some institutions are aware of the extent of the problem and are exploring changes to promotion criteria and mechanisms, redefinitions of academic roles, and so on. Effective mechanisms to recognise and reward excellent teaching have been well documented (Ramsden, 1995; Gibbs and Habeshaw, 2002). As efforts to balance teaching and research in appointment, probation and promotion decisions started in the UK 30 years ago (Gibbs and Openshaw, 1983), progress can sometimes seem painfully slow and easily derailed.

14. An increased emphasis on institutional 'USPs' and 'distinctiveness'

There is a good deal of debate, and controversy, about the value, and even the practicality, of institutions adopting a distinctive mission or approach to education, so as to attract a particular segment of the student market or improve a particular quality indicator. The GuildHE association of universities have made 'diversity and distinctiveness' their central value. The most

obvious claim to distinctiveness, given the use of quality indicators about employment and graduate salaries in KIS, is a clear focus on employability. Before the current HE market was introduced there were already institutions that had abandoned all 'academic' curricula and had adopted an entirely utilitarian and employment-oriented stance in their offer of programmes. In institutions without such an employability focus to curricula, employability agendas tend to involve extra-curricular initiatives. The USP may exist in the mission statement and in centrally run initiatives but not yet commonly in curricula or pedagogies.

Nevertheless there are signs that the kind of pedagogy that would be appropriate for improving employability is becoming better understood. Innovative practices are becoming more common, at least in individual departments if not yet commonly across institutions, exemplified in Pegg *et al.* (2012).

However, demand for traditional academic subjects that are not noticeably distinctive in content or pedagogy has not yet collapsed, perhaps in part because it is institutions with the highest reputations that offer such programmes.

Even in research universities with traditional curricula, these are now sometimes being audited by the centre to examine the extent to which the institutional strategic emphasis on employability is evident, or whether programme-level learning outcomes articulate with institutional statements about, for example, 'graduate attributes'. In some cases the form of both module-level and programme-level documentation has been changed so as to encourage all curricula to emphasise such graduate attributes. Changes in curricula of this kind take some years to be implemented fully and any impact on employability data would be hard to judge for some time. To see what the longer-term consequences of a comprehensive reworking of curricula around graduate attributes might be, both for establishing a distinctive market position and for an impact on employability, it is at present necessary to look at Australian universities, such as the University of Sydney.

There are examples of distinctive institutional pedagogies, implemented across all programmes, such as at The Open University and Oxbridge, but these are unlikely to be imitated because they are built into the infrastructure of the institutions to such an extent. Elsewhere in Europe there are examples of institutions with distinctive pedagogies that have evolved over time, such as problem-based learning at the University of Maastricht, and that span all subjects. However, these often started, very successfully, in one subject (such as Medicine, or Engineering, usually a professional subject) and were

gradually adapted to other disciplines with more or less success over an extended period of time. There tends to be a constant battle between disciplinary distinctiveness and institutional distinctiveness, with disciplines having more stamina.

Becoming institutionally distinctive is not a quick or easy process. There are a few examples of institutions embarking on a long and detailed programme of identifying current weaknesses and strengths, analysing the market, and redefining goals, negotiated carefully with staff, bringing everybody on board, and implementing widespread changes across the whole institution, driven by an understanding of organisational change. However, current efforts to become distinctive sometimes appear too rushed and too centrally driven to succeed and may not engage hearts and minds, or change values, to the extent that is necessary. Distinctiveness that is real rather than only claimed may reside in culture and ethos, such as in the way some small institutions care for their students, which are not easy to capture in quality indicators or even to communicate convincingly in prospectuses. Future developments may involve learning how to identify and communicate existing distinctiveness, where it exists, rather than in trying to manufacture it.

As students apply to study programmes, there may be more scope than is currently exploited in departments developing, and communicating, their own distinctiveness from other departments teaching the same subject at competing institutions (for example that they involve far more work-based learning or more sophisticated research training): important features of distinctiveness that may not be shared with other subjects in the same institution. Institutional claims that the whole institution is distinctive cannot easily encompass programme-level distinctiveness and may even mask it, contradict it, or constrain possibilities for programme-level distinctiveness. Institutional distinctiveness may also be less important to prospective students than is programme-level distinctiveness, except for a small minority of genuinely different institutions. It is easier to find examples in the literature of departments turning themselves around and becoming distinctive in a convincing and successful way than it is finding examples of such changes in whole institutions.

15. Rebuilding learning and teaching strategies

The long-running and well-funded HEFCE Institutional Learning and Teaching Strategy initiative was intended to go beyond what individual teaching enthusiasts, or individual teaching development projects, could achieve, and instead to emphasise the institutional infrastructure within which

teaching was developed: policies about requirements for initial training, promotion criteria, alignment of efforts across resource, library, estates and other institutional strategies, engaging departments with institutional goals, setting clear institutional priorities and goals and measuring their achievement, and so on. Evaluation of early institutional progress was very encouraging (Gibbs *et al.*, 2000) and institutional strategies have been well documented (e.g. HEFCE, 2001). These strategies, and the organisational frameworks that underpinned them, would have been an ideal foundation for efforts to improve NSS scores or to address other quality issues in the current market environment, or to pursue new USPs. For example, in some institutions departmental learning and teaching strategies were drawn up as part of an annual round of planning and funding, oriented towards institutional teaching priorities, so that, over time, institutional goals became embedded in departmental behaviour and so that quality enhancement efforts were real and devolved rather than centralised and only documented.

However, the focus of HEFCE's funding drifted away from strategic concerns and towards individual components of what ought to have been coherent strategies. The focus on locally appropriate institutional strategy was replaced by a focus on nationally specified tactics, and these tactics mainly concerned individual teachers and their training and accreditation, with all the limitations discussed above. The approach to institutional learning and teaching strategies in Wales was not funded and took a while to gain any purchase, and Scotland went down a different route, emphasising national themes for development (such as feedback) rather than institutional organisation and strategic approaches to institutional missions. The focus on coherent institutional strategy has to some extent been lost, and many institutions are now having to start again in building a coherent overall approach to the development of quality in teaching, in the new market environment.

Having reviewed a number of what institutions term their 'strategies for teaching and learning', as a part of the research for this publication, it is clear that many of the weaknesses observed 15 years ago are again evident. Strategies today sometimes appear to consist largely of a list of tactics, or stand-alone projects, targeted on specific short-term problems, with no coherent goals, no overall change strategy that would engage all departments or programmes, and no way of judging if it has been successful. This has been a very expensively lost opportunity.

16. Final comments

It is clear that collecting and publishing quality indicators about quality of degree programmes, and making these data available to prospective students in accessible forms, is galvanising higher education institutions to pay attention to a raft of teaching quality issues. So powerful is this effect that in many institutions it is in practice replacing previous quality assurance processes, changing quality enhancement mechanisms and leading to bold new initiatives designed to improve quality indicators and student experience.

What is less clear is whether this will always improve quality, particularly as some of the currently used quality indicators are invalid as predictors of learning gain and several important and valid quality indicators are missing from current data collection and publication, from quality assurance documentation and processes, and from universities' prospectuses. Prospective students are not yet provided with data about crucial aspects of educational provision they will receive in return for their fees (such as the volume of feedback and the size of classes) and have to rely largely on satisfaction ratings.

Institutions and departments appear to be paying attention to teaching quality as never before, and it may be that this attention, if intelligently focused, will improve quality whether or not it is targeted on improving current quality indicators and whether or not it actually impacts on current quality indicators. However, some efforts seem likely to compromise quality even if they improve the currently used quality indicators. For example, the large-scale rationalisation of course provision that is currently taking place seems likely to increase class sizes, even when the rationalisation is driven by quality indicators.

'Bureaucratic' change processes, in the sense that McNay (1995) uses the term, that rely on central policies and regulations, seem less likely to bring about worthwhile change than processes that allow scope to departments and subjects to address quality issues in locally appropriate ways, but give support to such local efforts and check that they take place and that they work.

Because current quality indicators are about programmes it seems essential that, where the current institutional focus of attention is on individual modules and individual teachers, this focus is shifted to programmes and programmes teams, both for quality assurance and quality enhancement. Career structures and recognition and reward mechanisms need to be adopted that support more effective programme leadership.

It is also unclear if prospective students actually use quality data to make choices about where to study, or about what level of fees it is worth spending on studying a programme. It will take further research on student behaviour and thinking to explore this issue.

Issues concerned with cost-effectiveness, value for money, and costing programmes in relation to the level of provision, seem not to have been addressed to any great extent by institutions, though this may change once the impact of fee income, and the level of student demand, becomes clearer. The relationship between fee levels and quality indicators is an obvious further area for future research.

17. Policy implications

The purpose of this report was to describe institutional behavior in a higher education market driven by quality indicators, and to suggest ways to make the use of indicators more effective. The analysis has significant policy implications.

- 1 Much of the institutional effort that this report describes is predicated on the assumption that students pay close attention to quality indicators when making decisions about where to spend their fees. There is limited evidence to support this assumption and this is an urgent area for research.
- 2 When the NSS and KIS are next reviewed there should be a tougher stance taken concerning the continued inclusion of variables that are not valid, in the sense that they do not predict learning outcomes or learning gains. In the case of the inclusion of 'class contact hours' in KIS, this should be removed as soon as is practicable in order to limit the damage it is causing through inappropriate changes in patterns of teaching that are likely to reduce learning gains even if they improve the quality indicator. It may be helpful to have deeper conversations with student organisations concerning which quality indicators are valid and useful, and which may have negative side effects.
- 3 It may be difficult to include more valid indicators of quality in KIS that are currently omitted, such as the amount of formative-only assessment and class and cohort size, due to the difficulty of defining standard quantitative indicators in ways that are easy to compute or to make sense of across very different contexts. Instead, prospectuses might be expected to include more information about these variables for each degree programme, in a way that is understandable to students in each context. This would give a fuller picture of what volume of provision students' fees

are paying for than is currently the case. It would make direct comparison between programmes more difficult but would be more informative. The National Survey of Student Engagement provides a more valid predictor of learning gains than does the NSS, as well as a clearer indication of the nature of provision that students experience, and current piloting of its use in the UK should be supported and extended.

- 4 Institutions should review their course documentation requirements concerning which educational variables should be included, to ensure that vital variables are not omitted and hence overlooked during course design, and so that courses are not approved that are characterised by features that are educationally undesirable (such as involving no formative-only assessment or no close contact with teachers).
- 5 National quality enhancement initiatives should concentrate more on whole degree programmes, and their design and development, and on the necessary associated leadership of teaching, and less on individual teachers, individual modules, and isolated innovations. The HEA's existing extension of teaching accreditation to more senior levels of academics could be accompanied by national recognition of outstanding leaders of teaching and outstanding degree programmes (to match or partially replace the existing national recognition of individual teachers). Institutional mechanisms to recognise and reward excellent teachers should shift their focus to excellent learning environments or degree programmes.
- 6 More evaluation could be undertaken concerning the usefulness of the NSS by relating NSS scores for degree programmes to other variables such as fee levels, student retention rates and the proportion of teachers who are trained and accredited. Ideally these would be multivariate analyses taking into account students' prior educational qualifications and other compounding variables.
- 7 NSS scores for a degree programme provide a useful indication of where there might be problems, but rarely point directly to the cause of problems let alone to possible solutions. It would be useful if, for each issue focused on by the NSS, such as feedback, there was a more detailed short questionnaire that institutions and departments could use themselves to explore local issues in more depth. There is also scope for offering guidance on the likely causes of, and possible solutions to, low scores on the different scales of the NSS, ideally linked to case studies of successful interventions to improve NSS scores. The NSSE provides more direct indications of what practices to pay attention to than does the NSS, and its pilot use in quality enhancement should be supported.

- 8 The ways in which institutions use NSS and other data to improve educational quality vary widely, and the effectiveness of different institutional strategies also appears to vary widely. This report has given a glimpse of some of the differences and some of the effective strategies, but there is plenty of scope for more detailed and varied case studies that could share effective institutional change strategies more widely. There are more than 500 web-accessible case studies of using the NSSE successfully to improve student engagement (NSSE, 2012). Institutions, and departments within subject associations, may balk at helping their competitors directly, but on the basis of the current study there seems no reluctance to make effective practices more public.
 - 9 A system that emphasises quantitative comparisons between units of analysis demands the use of standard definitions of what those units consist of. A side effect appears to be a loss of curriculum diversity, including of multi- and inter-disciplinary degree programmes and modular programmes that allow mixing of subjects within a degree, and do not correspond well to the conventional subject labels used for comparison. The scale of this side effect, and its impact on innovation in curricula and on student choice and engagement, may need to be monitored³.
 - 10 As institutions attempt to become more distinctive and to carve out market niches, it will be important to identify if these identities are reflected in improved, or comparatively strong, performance in the emphasised areas: in other words whether there is any substance to the claims of distinctiveness. To do this, it may be necessary to collate other quality indicators than those currently used. In particular current quality indicators concerning employability may not discriminate between contexts in ways that are interpretable.
 - 11 A future review of the purpose of national quality assurance mechanisms, and their operation, should take into account the reality that quality assurance and enhancement within institutions is increasingly dominated by annual comparative review of quantitative data about educational provision, rather than by periodic qualitative review of quality procedures.
 - 12 There should be a national-scale initiative that provides evidence-informed guidance, and shares effective practice, concerning cost-effective teaching practices, so that where reductions in provision are forced by reductions in funding, these are less likely to cause reductions in learning.
- 3 The HEA has identified the impact of recent changes in HE on learning and teaching as its overarching research and policy theme for the next three years.

References

- Astin, A.W. (1985) *Achieving educational excellence: a critical assessment of priorities and practices in higher education*
- Lord Browne et.al. (2010) *Securing a sustainable future for Higher Education: an independent review of Higher Education Funding & Student Finance*
- Chickering, A.W. and Gamson, Z.F. (1987) *Seven Principles for Good Practice in Undergraduate Education*. Racine, WI: The Johnson Foundation Inc.
- Coffey, M. and Gibbs, G. (2000) The evaluation of the Student Evaluation of Educational Quality questionnaire in UK higher education. *Assessment and Evaluation in Higher Education*. 26 (1), 89-93.
- de Bruijn, H. (2002) Performance measurement in the public sector: strategies to cope with the risks of performance measurement. *International Journal of Public Sector Management*. 15 (7), 578-594.
- Dunne, E. and Zandstra, R. (2011) *Students as change agents*. Bristol: ESCalate.
- Ehrenberg, R.G. (2006) *What's Happening in Public Higher Education?* Westport, CT: Praeger.
- Entwistle, N.J. and Tait, H. (1994) *The Revised Approaches to Studying Inventory*. Edinburgh: Centre for Research into Learning and Instruction, University of Edinburgh.
- Freeman, T. (2002) Using performance indicators to improve health care quality in the public sector: a review of the literature. *Health Services Management Research*. 15 (2), 126-137.
- Gibbs, G. (1995) National scale faculty development for teaching large classes. In: Wright, A. (ed.) *Teaching Improvement Practices*. New York: Anker.
- Gibbs, G. (2010) *Dimensions of quality*. York: Higher Education Academy.
- Gibbs, G. (2013) Reflections on the changing nature of educational development. Invited commentary. *International Journal for Academic Development*. 18 (1). Forthcoming.
- Gibbs, G. and Coffey, M. (2004) The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. *Active Learning in Higher Education*. 5 (1), 87-100.

- Gibbs, G. and Dunbar-Goddet, H. (2009) Characterising programme-level assessment environments that support learning. *Assessment and Evaluation in Higher Education*. 34 (4), 481-489.
- Gibbs, G. and Habeshaw, T. (2002) *Recognising and Rewarding Excellent Teachers*. Milton Keynes: Open University, National Co-ordination Team for the Teaching Quality Enhancement Fund.
- Gibbs, G. and Openshaw, D. (1983) *Rewarding Excellent Teachers*. SCEDSIP Paper No. 14. Birmingham: SCEDSIP.
- Gibbs, G., Habeshaw, T. and Yorke, M. (2000) Institutional learning and teaching strategies in English higher education. *Higher Education*. 40 (3), 351-372.
- Gibbs, G., Regan, P. and Simpson, O. (2006) Improving student retention through evidence based proactive systems at the Open University (UK). *College Student Retention*. 8 (3), 359-376.
- Graham, A and Thompson, N. (2001) Broken ranks: US News' college rankings measure everything but what matters. And most universities don't seem mind. *Washington Monthly*. 33 (4), pp9-14
- Hattie, J.A.C. (2009) *Visible Learning: a synthesis of meta-analyses relating to achievement*. Abingdon: Routledge.
- HEFCE (2001) *Analysis of Strategies for Learning and Teaching*. Report 01/37a. Bristol: Higher Education Funding Council for England.
- HEPI (2007) *The Academic Experience of Students in English Universities (2007 Report)*. Oxford: Higher Education Policy Institute.
- Jessop, T., McNab, N. and Gubby, L. (2012) Mind the Gap: An analysis of how quality assurance processes influence programme assessment patterns. *Active Learning in Higher Education*. 13 (2), 143-154.
- McNay, I. (1995) From the collegial academy to the corporate enterprise: the changing culture of universities. In: Schuller, T. (ed.) *The changing University?* Buckingham: SRHE and Open University Press.
- Marsh, H.W. (1982) SEEQ: a reliable, valid and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology*. 52 (1), 77-95.

Nasr, A., Gillett, M. and Booth, E. (1996) Lecturers' teaching qualifications and their teaching performance. *Research and Development in Higher Education*. 18, 576-581.

National Survey of Student Engagement (2007) *Experiences that matter: Enhancing student learning and success*. Bloomington, IN: Indiana University Center for Postsecondary

National Survey of Student Engagement (2012) *How Institutions Use NSSE*. Available from: http://nsse.iub.edu/_/cid=315 [October 2012].

Pascarella, E.T. and Terenzini, P. (2005) *How college affects students: a third decade of research. Volume 2*. San Francisco: Jossey-Bass.

Pascarella, E.T., Seifert, T.A. and Blaich, C. (2010) How effective are the NSSE benchmarks in predicting important educational outcomes? *Change: The Magazine of Higher Learning*. 42 (1), 16-22.

Pegg, A., Waldock, J., Hendy-Isaac, S. and Lawton, R. (2012) *Pedagogy for employability*. York: Higher Education Academy.

Propper, C. and Wilson, D. (2003) The Use and Usefulness of Performance Measures in the Public Sector. *Oxford Review of Economic Policy*. 19 (2), 250-267.

Ramsden, P. (1995) *Recognising and Rewarding Good Teaching in Australian Higher Education*. Canberra: Australian Government Publishing Services.

Ramsden, P. (1999) A performance indicator of teaching quality in higher education: the Course Experience Questionnaire. *Studies in Higher Education*. 16 (2), 129-150.

Smith, P. (1990) On the unintended consequences of publishing performance data in the public sector. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*. 153 (1), 53-72.

Twigg, C.A. (2003) Improving Quality and Reducing Cost: Designs for Effective Learning. *Change: The Magazine of Higher Learning*. 35 (4), 22-29.

Yorke, M., Bridges, P. and Woolf, H. (2000) Mark distributions and marking practices in UK higher education; some challenging issues. *Active Learning in Higher Education*. 1 (1), 7-27.

Contact us

The Higher Education Academy
Innovation Way
York Science Park
Heslington
York
YO10 5BR

+44 (0)1904 717500
enquiries@heacademy.ac.uk

ISBN: 978-1-907207-62-4

© The Higher Education Academy, 2012

The Higher Education Academy (HEA) is a national body for learning and teaching in higher education. We work with universities and other higher education providers to help bring about change in learning and teaching. We do this to improve the experience that students have while they are studying, and to support and develop those who teach them.

Our activities focus on rewarding and recognising excellence in teaching, bringing together people and resources to research and share best practice, and by helping to influence, shape and implement policy - locally, nationally, and internationally.

www.heacademy.ac.uk | Twitter @HEAcademy

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any storage and retrieval system without the written permission of the Editor. Such permission will normally be granted for educational purposes provided that due acknowledgement is given.

To request copies of this report in large print or in a different format, please contact the communications office at the Higher Education Academy: 01904 717500 or pressoffice@heacademy.ac.uk

The views expressed in this publication are those of the authors and not necessarily those of the Higher Education Academy.