Teaching Quality in Higher Education: Literature Review and Qualitative Research
May 2016

Prepared for the Department for Business, Innovation and Skills by HOST Policy Research.

Authors: Professor David Greatbatch and Dr Jane Holland

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Business, Innovation and Skills.

Executive Summary

Introduction

Since the end of the 1980s higher education (HE) in the UK has been re-shaped, mainly in response to pressures emanating from socio-economic change in society, the opening up of HE to the market through the introduction of student fees, government policies aimed at widening participation, and the application of Computer and Information Technology. The HE response has resulted in a fundamental change - the move from an elite system, to a mass system to which at least 40% of the population has access. The 'market' subjects HE to the needs and aspirations of society and this notion has been incorporated into government policies administered through the funding councils, along with the introduction of student tuition fees to finance undergraduate courses.

It is against this background that the government intends to introduce the Teaching Excellence Framework (TEF). The aims of TEF are to: recognise and reward high quality teaching; provide prospective students with information that will help inform their decisions about which institution to attend; and assist employers, who currently rely on proxy measures for teaching quality that are highly problematic (such as research success and institutional reputation) in their recruitment. This study is therefore extremely timely, as it offers an overview and critical evaluation of the latest evidence on definitions, initiatives, metrics and the operationalisation of teaching excellence initiatives in HE, both nationally and internationally.

Objectives of the project

HOST Policy Research (HOST) was appointed by the Department for Business, Innovation and Skills (BIS) to conduct a literature review and qualitative research exploring teaching quality in higher education. The headline aims of the project were to ensure that BIS has access to the most up-to-date research and current thinking in this area, with a good understanding of the context of any potential issues relating to assessing Teaching Excellence in HE. The specific objectives of the project were to:

- Review the literature on attempts to measure Quality across HE in the UK and internationally, defining 'Excellence' as an element of 'Quality' and potential metrics for measurements of teaching quality.

- Consult experts in the field to ensure the literature review is comprehensive and to identify unpublished work.

- Qualitatively assess student opinions of teaching quality, in order to gain a better understanding of what informs these.

Methodology

Stage 1: Literature review Stage

2: Qualitative Research

- Step 1: Consultations with experts
- Step 2: Focus groups with students

Findings from the literature review

Attempts to measure teaching quality across HE

Overall, the literature suggests that external assessments of teaching quality have stimulated increased attention to the quality of teaching in UK higher education. However, studies of previous teaching quality regimes in the UK imply that assessments were influenced by universities' research reputations and other factors that were not directly linked to teaching quality. Some studies argue that teaching quality audits inhibited innovation in approaches to teaching because they diverted academic staff time towards administrative tasks rather than seeking to bring about a change. It is also felt that conceptions of teaching quality vary across different disciplines/subjects - though this issue is not thoroughly explored in the literature.

International attempts to measure quality across HE

Governments in many countries have set up quality assurance agencies to conduct external audits (including assessment of teaching and learning) of higher education institutions. Currently, there is limited research on the extent to which such audits have improved quality assurance in universities. However, studies suggest that external quality audits, together with internal university processes, have stimulated change and improvement of teaching and learning in some countries. Research shows the importance of considering the 'micro processes' surrounding teaching and learning, such as lecturers' teaching strategies and the characteristics of university students' learning, however, there is very little research on the impact of audits by external agencies on these. There is an emerging interest in how students in different disciplines have differing expectations from their courses but it remains unclear whether and how this impacts on their perceptions of teaching quality and course evaluations.

Research on quality assurance systems designed to measure teaching quality across HE in China and Germany points to the importance of a three-stage process involving self- evaluation by HEIs (stage one), external evaluation by peer reviewers (stage two), and implementation of peer reviewers' recommendations (stage 3). In China, the ranking of the overall teaching quality of each institution as excellent, good, pass or unqualified during the first five-year cycle of evaluations of teaching quality across HE (2003-2008) was widely challenged by scholars due the high number of HEIs that were ranked as excellent. The first cycle of evaluations of teaching quality was also criticized for failing to take students' learning outcomes into account. As a result, the second cycle of evaluations (2012-ongoing) is placing a greater emphasis on actual teaching processes in HEIs, and the measures of quality are being modified to incorporate student learning.

The study of systems used to evaluate teaching quality in HEIs in Germany reveals that they enjoyed wide acceptance and were seen to be useful and effective. However concerns were raised about the need to: reduce the work burden associated with the systems; broaden the composition of expert panels (to include, for example, students, international experts, and people working in professional practice); and ensure that reviewers' recommendations were more evenly distributed between 'planning and organization of teaching and learning' and 'resources' (which were common) and 'goals for teaching and learning'' (which were rare).

Research on the Australian Learning and Teaching Council's criteria for excellence in university teaching indicates that criteria and metrics used to identify and measure effective teaching need to be continually reviewed in order to ensure that they reflect changing definitions of effective teaching due to changing student needs, new modes of delivery, changing employer needs and technological innovation.

Teaching Excellence in HE

European Association for Quality Assurance in Higher Education's (ENQA) report on the Concept of Excellence in Higher Education reports that Quality Assurance systems for Higher Education across the world are not particularly geared to identifying or promoting excellence, but are more about setting thresholds for quality and encouraging improvement.

Defining excellence is considered to be a challenge and the literature provides a plethora of definitions of excellence. It is a contested concept, with varied competing interpretations**.** The situation is complicated by the fact that excellent teaching is not only down to the individual teachers and the learners they work with, but also the supporting resources and systems, which facilitate the learning process. Excellence can also be interpreted using either norm-referenced definitions of excellence, which define excellence relative to the performance of others and imply competition and elitism, or criterion- referenced definitions of excellence, which define excellence in relation to a standard such that people are not in competition for it.

There are few empirical studies on academics' views about teaching excellence and what they regard as good indicators of excellent teaching, perhaps because of the difficulties associated with defining the term precisely. However, studies of teaching excellence as perceived by students indicate that students want to be taught by staff who are enthusiastic and knowledgeable about their subject, empathetic, approachable, helpful and patient, and encourage students to develop their full potential.

Metrics for measuring teaching quality

Systems of measurement (where discussed) tend to draw on a basket of measures, due to the wide range of potential metrics. Quantitative metrics can be grouped into measures of inputs and outputs, process measures and rankings. Many authors are dubious about the value of key input (prices, fees, staffing numbers) and output measures (degrees awarded, jobs following graduation, raw student satisfaction scores).

There is more support for process measures looking at student engagement (though not systematically studied), but without clearly tried and tested tools for use in England. There

is some evidence that students favour student engagement surveys over satisfaction surveys, and may take them more seriously.

Ranking metrics are generally regarded with caution, as the most respected rankings are based more on research excellence rather than any consideration of teaching excellence. There is little evidence in the literature on qualitative metrics, although a number of authors comment on the desirability of quantitative metrics of teaching quality being supported by some form of peer review, accreditation, visit or audit.

Findings from student focus groups

Teaching quality did not feature as a key issue that the students explicitly considered when applying to university. However this is not to say that would have been the case had more information been available- indeed, some students appeared to have used the reputations of universities, departments and academic staff as proxies for teaching quality. Issues students considered included university reputation, reputation of department; research profile of the department; and whether an institution was within commuting distance. Other factors mentioned were course content, the general 'feel' of the university ("can you see yourself living here?"), location, personal experiences and observations at open days and interviews, universities prospectuses online, other websites – such as Which University? – which provided information on universities, league tables which rank universities, employment outcome information on the UCAS website, and social media. For many students, fellow students' comments about their experiences at universities on social media websites such as Facebook and The Student Room were the most useful and trusted sources of information..

Most students were agreed that they would have considered TEF data when applying to university (apart from those who were set on applying to a particular institution), but many doubted whether TEF would have influenced their decision. The students were, in general, not very enthusiastic about a TEF at the institution level, although there were those who thought it might be of some value. Some students in post-1992 universities were sceptical of the value of such ratings because they felt "they always favour elite institutions". It was suggested that the TEF should be conducted at subject level due to differences in approaches to teaching in different disciplines/subjects. Some students were sceptical about the inclusion of employment outcomes within the TEF because they felt "graduate employability is not a measure of teaching quality; however, some thought data on the extent to which graduates secure appropriate employment after their degree would be useful (whether as part of the TEF, or alongside the TEF). Some students felt that the value of a TEF would be felt more by employers in an 'increasingly competitive [market] for graduate jobs.

Whilst most students were generally happy with both the overall quality and value for money of their programmes, some were not satisfied, citing lack of resources, poor lecturing styles, unhelpful or untimely feedback, and too few contact hours with appropriate staff. The latter are distributed across all of the institutions involved in the study.

Lecture and seminar class/tutor group sizes were a common concern, however. Interestingly, a large majority of students said that they were academically stretched and

challenged and were engaged and motivated to pursue self-study. Students largely felt that their courses were helping them develop the knowledge and skills they need for the future. This was especially so in the case of students whose courses involved placements linked to employment after graduation. However, some felt that the careers services and provision of career guidance could be more visible, especially during the first two years of degree programmes.

Issues to consider

The lack of agreement in the literature as to what constitutes teaching excellence in higher education is striking. Comparative research looking at perceptions of teaching excellence across different disciplines, universities and countries is needed. However, this is unlikely to be an easy task, as just as definitions of quality change over time and from one context to another, so too do definitions of excellence.

Future research on teaching quality in HE should take on board the complex relationships between proxies for teaching quality and actual teaching quality. Particular attention should be paid to how definitions of effective teaching evolve over time due to changing student needs, new modes of delivery, changing employer needs and technological innovation.There is relatively little research into academics' understanding and perceptions of teaching excellence and further research in this area is sorely needed. It would be particularly helpful to look at how academics and students from the same institutions and courses perceive teaching quality and excellence, and also how these perceptions may change over time. There is also the issue of shifting thresholds, and how these can be accommodated in systems designed to measure excellence.

There is a relative paucity of literature on the potential for using types of learning engagement as a metric in a set aimed at identifying learning excellence. This may warrant further detailed investigation as a separate research topic when developing a potential battery of metrics for teaching excellence.

There is no evidence on how teaching quality metrics can be broken down to different groups of learners, particularly groups differentiated by learner characteristics rather than subject/programme studied. The metrics that might pick that up – student engagement – are measured on voluntary completion of survey instruments, which may not be applied to all final year undergraduates in an institution, and will certainly not be completed by all final year undergraduates. Survey instruments are in any case meant to be anonymous. Disadvantage and under-representation will also be specific to programme and to institution, depending on catchment, fees and other characteristics.

Neither the benefits accruing to HEIs from an increase in teaching quality nor the views of employers regarding teaching quality in HE have been systematically explored in the UK. Research is urgently required in both of these areas.

Forthcoming Research

A rich body of research is emerging in relation to the Undergraduate Teaching Evaluation (UTE scheme in China, which provides insights that may be relevant in England, especially in relation to the use of student engagement and course evaluation surveys in national HE teaching quality assessment systems.

Contents

Executive Summary **3**

1. Introduction **10**

1.1 Understanding the environment 10

1.2 Objectives of the project 11

1.3 Methodology 12

1.4 Structure of the report 13

2. Measure teaching quality **14**

2.1 Introduction 14

2.2 Teaching Quality Assessment 14

2.3 Subject Review 15

2.4 Institutional Audit 16

2.5 Summary 17

3. International Models **18**

3.1 Introduction 18

3.2 External audits of higher education institutions 18

3.3 Measuring teaching quality across HE 19

3.4 Summary 24

4. Teaching excellence in HE **26**

4.1 Introduction 26

4.2 Defining excellence 26

4.3 Perceptions of excellence 28

4.3 Measuring Excellence 30

4.4 Summary 31

5. Metrics **32**

5.1 Introduction 32

5.2 Quantitative metrics 33

5.3 Qualitative metrics 40

5.4 Summary 40

6. Student Focus Groups **41**

6.1 Introduction 41

6.2 Applying to university 41

6.3 How TEF may affect the students' behaviour and decisions 45

6.4 Current views on teaching quality within HE 48

6.5 Summary 52

7. Issues to consider..... 53

References..... 54

Annex 1 - Metrics: Strengths and Weaknesses 66

Annex 2 - Focus Group Topic Guide..... 71

1. Introduction

1.1 Understanding the environment

Since the end of the 1980s higher education (HE) in the UK has been re-shaped, mainly in response to pressures emanating from socio-economic change in society, the opening up of HE to the market through the introduction of student fees, government policies aimed at widening participation, and the application of Computer and Information Technology. The HE response has resulted in a fundamental change - the move from an elite system, to a mass system to which at least 40% of the population has access. Consequently, the student population has become increasingly diverse in terms of: age on entry; previous experiences; qualifications; socio-economic status; ethnicity; and cultural background. This has been facilitated by: the creation of new routes into HE (e.g. access courses and Foundation Degrees); modularisation of the HE curriculum; re-structuring the academic year into semesters; and the redesigning of courses at entry level to ease the students' transition to HE.

In the light of the changes taking place in HE and the emergence of the view of students as 'consumers' with rights and entitlements (Barnett 2011; McCulloch, A. 2009; Department for Business, Innovation & Skills 2011), concerns have arisen about the quality of teaching and learning and the provision of student support. In 1997, Dearing charged the HE sector with the pursuit of 'excellence' in 'teaching and learning' (Dearing1997). As a result, for over two decades curriculum change has been driven by a range of organisational agents and by the funding of development programmes, together with an increasing emphasis on professionalising teaching in HE by challenging the assumption that whilst university staff may be specialists in their subjects they are not necessarily experts in the practice of education. The response of the HE system to these initiatives has included: a shift to learning outcomes models; the stipulation of aims and objectives in module and programme specifications; an increase in subject benchmarking; assessment becoming increasingly complex and seen as a key driver for learning; and a greater emphasis on staff development. In addition, the external forces for change to which the HE sector has been subjected have been accompanied by those which are self- determined from within, including the commitment of academic staff to enhancing the curriculum by improving their professional practice with regard to the design and delivery of teaching and learning.

The 'market' subjects HE to the needs and aspirations of society and this notion has been incorporated into government policies administered through the funding councils, along with the introduction of student tuition fees to finance undergraduate courses. Following the introduction of student tuition fees, the expectation is that entrants to HE will not only be seeking a university education, but value for money as well (Foskett, Roberts and Maringe 2006; Barnett 2011). This has led to a growing view, reflected in government policy, that students are becoming, 'choosier and more demanding consumers of the higher education experience' (Mandelson 2009). The overall effect of this has been to move teaching and learning in HE in the direction of greater accountability and responsiveness to student choices and expressions of their satisfaction (or otherwise).

It is against this background that the government intends to introduce the Teaching Excellence Framework (TEF). The aims of TEF are to: recognise and reward high quality

teaching; provide prospective students with information that will help inform their decisions about which institution to attend; and assist employers, who currently rely on proxy measures for teaching quality that are highly problematic (such as research success and institutional reputation) in their recruitment. This is far from a straightforward exercise because the concepts of teaching quality and teaching excellence and the use of metrics for measuring them are contested areas (Land and Gordon 2015). A recent review of the research and grey literature on teaching excellence and quality (Gunn and Fisk 2013) highlighted a number of issues, including: a lack of articulation around the differences between threshold quality and teaching excellence; a lack of sophistication in conceptualisation of university teaching excellence; a lack of representatively diverse conceptualisations of teaching excellence, which suggests a normative universalising of teaching excellence; and a significant gap between teaching excellence as practice and educational theory concerning teaching excellence. This study is therefore extremely timely, as it offers an overview and critical evaluation of the latest evidence on definitions, initiatives, metrics and the operationalisation of teaching excellence initiatives in HE, both nationally and internationally.

1.2 Objectives of the project

HOST Policy Research (HOST) was appointed by the Department for Business, Innovation and Skills (BIS) to conduct a literature review and qualitative research exploring teaching quality in higher education. The headline aims of the project were to ensure that BIS has access to the most up-to-date research and current thinking in this area, with a good understanding of the context of any potential issues relating to assessing Teaching Excellence in HE. The specific objectives of the project were to:

- Review the literature on attempts to measure Quality across HE in the UK and internationally, defining 'Excellence' as an element of 'Quality' and potential metrics for measurements of teaching quality

- Consult experts in the field to ensure the literature review is comprehensive and to identify unpublished work.

- Qualitatively assess student opinions of teaching quality, in order to gain a better understanding of what informs these.

The project was designed to help interpret some findings from the new Higher Education Green Paper and inform the technical consultation on the operational details of the Teaching Excellence Framework (TEF).

1.3 Methodology

The methodology had two core elements: literature review and qualitative research. The methodology is summarised below.

Phase 1: Literature Review

The review of the literature on teaching quality in higher education covered both academic literature and the grey literature. The literature review was undertaken systematically, using a protocol-based approach to search for, and retrieve, relevant literature based on pre-specified inclusion and exclusion criteria agreed with BIS. These materials were then summarised, synthesised and critically evaluated. The literature review involved six stages:

- Development of a research protocol detailing the steps and procedures to be followed in the literature review including: the locations/sources to be searched for literature; the search terms to be used to retrieve the literature; the screens each item would pass through in being considered for inclusion in the review (the inclusion and exclusion criteria); processes for recording and storing references; and processes for retrieving and summarising literature.
- Searches of a wide range of online databases and websites, which offer electronic access to most published literature in the area of teaching quality in higher education in England and internationally.
- Screening literature for inclusion in the review using the following inclusion criteria: academic literature (journal articles, books, book chapters), including both empirical and conceptual studies; grey literature (information or research output produced by organisations or individuals, outside of commercial or academic publishing and distribution channels, such as institutional or technical reports, working papers, conference proceedings, conference literature, and monographs, specialist literature, and primary data sources); editorials, newspaper articles and other forms of popular media which provide relevant information and evidence.
- Assessment of the quality of potentially eligible articles, books and reports (screening for exclusion) to ensure that the best available evidence was used in the review.
- Systematic extraction of data and evidence using templates for each of the countries included in the review, with headings that match the key research questions.
- Synthesis and critical evaluation of the data

Phase 2: Qualitative Research

Step 1: Consultations with experts: The purpose of these consultations was to: ensure that the literature review is comprehensive; and collect additional information in areas in which there are gaps in the literature. Potential interviewees were selected according to their contributions to and/or knowledge of the literatures on teaching quality and teaching excellence in HE. Telephone and email discussions were conducted following an agreed topic guide.

Step 2: Focus groups with students: We conducted focus groups at 10 higher education institutions. We sought to secure the involvement of students from a diverse range of backgrounds and disciplines in both pre- and post-1992 universities. The focus group discussion guide was developed in close consultation with BIS. The guide covered students' views on teaching quality and the proposed Teaching Excellence Framework. The focus group responses were written up into report templates to facilitate their analysis and integration into the wider project.

Reporting and outputs:

During the project we delivered fortnightly progress reports. A draft final report was discussed with the BIS steering group and, following receipt of comments, this final version was prepared.

1.4 Structure of the report

The report structure is as follows: Section 2:

Measuring teaching quality

Section 3: International Models

Section 4: Teaching Excellence in higher education

Section 5: Metrics

Section 6: Qualitative research

Section 7: Issues to consider

References

2. Measure teaching quality

2.1 Introduction

In the UK, teaching quality in higher education has been a topic of wide ranging debate (e.g. Dearing 1997, Macfarlane 2011; Department of Business Innovation and Skills 2011; and the recent Government Green Paper). Despite this, many questions remain to be answered in relation to how 'quality' can be evaluated across the HE sector. In this section, we examine these questions through an overview of the literature on previous national teaching quality regimes in the UK.

2.2 Teaching Quality Assessment

Since 1992, a series of national teaching quality regimes have been introduced in England starting with what became known as the Teaching Quality Assessment (TQA), which operated from 1993 to 1997. TQA provided external review and graded judgments of the quality of teaching and learning at institutional level (Cheng 2010). The TQA allowed institutions to select their own aims, adopted a broad understanding of teaching, based on the management and promotion of student learning, and involved students in the assessment of the quality of the teaching they were receiving. It involved a system based on self-assessment, external assessment, and peer review (Taggart 2004), with three possible assessment outcomes - "unsatisfactory, satisfactory and excellent" (Laughton. 2003).

TQA proved to be controversial with critics pointing to biases and weaknesses within the assessment system and calling for a greater degree of transparency in the process (Alderman 1996; Drennan and Beck 2001). Barnett (1994) found that the list of institutions achieving 'excellent' scores in the teaching quality assessments were highly correlated with the established elite universities in the UK, that is, the traditional 'old' universities. This led him to suggest that a cross-subject set of criteria was operating, which favoured certain kinds of institution, despite the Higher Education Funding Council for England's (HEFCE) intention to be mission sensitive. Drennan (1998) came to a similar conclusion in her study of the TQA scores of the 13 Scottish universities between1993 and 1998. Drennan examined the relationships between TQA scores and such factors as university age, research output as measured by the Research Assessment Exercise, entry point age of students, staff student ratio and library spend. She found a strong correlation between TQA results and the age of the university, with 4 ancient universities generally obtaining the highest scores, followed by the 4 moderns and 5 post-1992 universities. Drennan concluded that:

> "Rather than being an objective and value-free process, the TQA exercise brings a subjective focus into play. Assessments are less likely to be mission-sensitive and more likely to be influenced by pre-existing research reputations, and levels of faculty, which favour the older universities and disadvantage the new".

An evaluation of TQA commissioned by HEFCE (Barnett et al. 1994) proposed a range of changes including: universal visiting, summative judgement at the threshold level, and the

framing of recommendations for improvement within a limited number of dimensions, so as to produce a 'quasi-profile'.

2.3 Subject Review

Following the formation of the Quality Assurance Agency for Higher Education (QAA), TQA was replaced by Subject Review, which operated from 1998 to 2001. Subject Review implemented most of the changes proposed by Barnett et al. (1994) and was significantly different from that used previously (Laugnton 2003; Cheng 2010).

QAA's Subject Review focused on assessing teaching at subject levels. Its methodology involved review against aims and objectives at subject level and a narrative of self- evaluation, and the grading (by academic peers) of a number of components which constituted an overall profile. Six aspects were identified as a focus for reviewers, and each could score up to a maximum of four (making an important and substantial contribution to the aims and objectives of the provider, with all of these being met): curriculum development; content and organisation; teaching, learning and assessment; learning resources; student progression and achievement; student support and guidance; and quality and standards management and enhancement (Laughton. 2003). The process normally involved a four-day visit during which teaching was observed, student work was examined and documentation, of both the subject area and institutional quality assurance practices, was reviewed. The result was a numerical score for each aspect of provision on a four-point scale. A report of the outcomes of each institutional subject review was published and made available for public scrutiny (Gosling and Andrea 2001). The newspapers used Subject Review scores for their institutional rankings, initially against QAA's intentions (Brown 2013).

The Subject Review methodology also proved to be controversial, however. There was strong criticism of the Subject Review's tight management of the university sector, especially in relation to the work of academics (Cheng 2010), the administrative burden associated with the review system (Kogan, Bauer, Bleilie, & Henkel, 2000; Laughton 2003; Harvey 2005; Brown 2013) and the rising costs of meeting QAA requirements (e.g. Harrison and Lockwood 2001). Several empirical studies of the Subject Review (and its predecessor the TQA) suggested that the audit diverted academic staff time towards administrative jobs and thereby actually undermined the quality of teaching and learning (Harvey 2005; Morley 2003; Newton 2000). Concerns were raised that some academics would become compliant with the audit rather than seeking to bring about a change in the learner or practice a diversity of approaches to teaching (Harvey, 2010; Srikanthan & Dalrymple, 2003). The compliance culture it was argued dampened creativity and slowed down the responsiveness of the university to a rapidly changing environment (Harvey & Newton, 2007; Henkel, 2000; Kogan, Bauer, Bleilie, & Henkel, 2000). Some studies suggested that only limited improvements had been made in relation to student experience and engagement in learning despite the growth of quality audits (Harvey 2002; Newton 2000 and Shah 2012-763-764).

Studies of the period when the Subject Review was operating suggest that quality monitoring in the UK was beset by overlapping and burdensome processes (Harvey 2005; see also Kogan et al. 2000; Laughton 2003), as well as competing notions of quality, a failure to engage learning and transformation, and a focus on accountability and compliance (Shah 2012-763-764). A study at Sheffield Hallam University, which sought

academics' perceptions of the different external monitoring processes in England, found that some processes, notably external examination, were perceived as providing a check on standards, however, there was little support for the view that external quality evaluation improved the student learning experience (Harvey and Newton 2004). Concerns were also raised that a rising trend of grades through time was due to institutions learning how to write self–created aims and objectives so as to maximize their grades (*THES*, 2 March 2001, cited in Harrison and Lockwood 2001; see also Laughton 2003).

Once again, as with the TQA, empirical studies provided evidence of an elitist bias within the system, which implicitly and structurally advantaged relatively resource rich older universities (Laughton 2003: 314). Drennan and Beck (2001) examined the relationship between Subject Review assessment scores, reputational factors and resourcing indicators. They found that from 1993 to 2000 "the determining factors in relation to high TQA scores (were) age of institution and reputation, in terms of attractiveness to students with high entry qualifications, and in terms of peer-assessed research. These variables (were), in turn, strongly related to other variables such as the staff-student ratio and library and computer spending" (Drennan and Beck 2001). Cook's (2001) analysis of 665 QAA reports from the 1998 to 2000 inspection rounds revealed "the richest, old universities outperforming their poorer counterparts. In 1998–2000, the average mark among old universities was 22.33 compared with 21.58 among post 1992 universities and 20.5 among colleges with higher education provision" (THES 2 March, 2001, p. 5). These studies lent support to the view that rather than representing an independent indicator of teaching quality, external quality review scores had largely served to confirm existing hierarchies, which might have been determined by factors other than teaching (Drennan and Beck 2001; see also: Laughton 2003).

Cook, Butcher and Raeside (2006) subsequently examined the grades awarded in the subject level review process that ran in England between 1995 and 2001. They noted that

> "These grades have subsequently been used to sustain most published league tables and are often cited in debates about institutional quality. However, the grades were never subject to any moderation, either at the time of the individual visit or subsequently. Accordingly, there are substantial variations between subjects and marked grade inflation over time" (Cook et al. 2006)

Their paper considered how to control for these effects and, having done this, what the impact on the published league tables would be. The difference between the adjusted and unadjusted scores were significant and, they concluded, further called into doubt both the validity of the original process and the uses to which the outputs had subsequently been put. Cook et al saw this as a warning to other quality assurance agencies which had used or are were considering using audits based on subject review.

2.4 Institutional Audit

Following significant pressure from HEIs for change, in 2001/2 the Subject Review was replaced by the Institutional Audit of the QAA, subsequent variations of which have been referred to Institutional Review from 2011 and Higher Education Review from 2013/14. The Audit is perceived as 'light touch' management to ensure that a higher education institution's quality systems are working effectively to maintain quality and academic standards (Harvey, 2005). Conducted at institutional level, the institutional audit focuses

less on directly assessing academic teaching at the subject level, and more on ensuring whether an institution's quality assurance systems are working effectively to maintain academic standards and quality (Harvey 2005). Laughton (2003) argued that the decision to pursue a 'light touch' methodology (and shift from subject level to audit at the institutional level) and, as a consequence, to abandon the phased introduction of a modified system of Subject Review known as Academic Review was a crucial turning point in the evolution of the quality assurance system of HE" in the UK:

> "Academic Review was intended to venture into an area of academic life which Subject Review confronted only tangentially - the issue of standards. As such, it had potential for being equally as controversial as Subject Review. However, this issue was moved from centre stage by the decision to pursue a modified approach based on the principle of 'light touch' which engendered a quality assurance framework containing a mixture of audit and review principles and practices." (Laughton 2003: 319)

Since the introduction of the institutional audit, a growing "emphasis" on information for students, student-based funding and student choice as drivers of quality, has led to concerns about more effort being put into things such as the National Student Survey, the Key Information Set and the commercial league tables, rather than quality enhancement (Brown 2013: 434). In addition, academics continue to express concerns about the bureaucratic demands which are seen, amongst other things, as having a detrimental impact on the activities of teaching staff (Feigenbaum and Iqani 2015).

Westerheijden, Hulpiau, and Waeytens (2007) contend that the views of academics are crucial because the impact of both external and internal quality assurance is dependent on them buying into the process. Academics' support is thus seen "as (…) a crucial factor influencing the accuracy and meaningfulness of the results achieved" (Cardoso, Rosa, and Santos 2013; Cardoso et al 2015; Laughton 2003).

2.5 Summary

Overall, the literature suggests that external assessments of teaching quality have stimulated increased attention to the quality of teaching in UK higher education. However, studies of previous teaching quality regimes in the UK imply that assessments were influenced by universities' research reputations and other factors that were not directly linked to teaching quality. Some studies argue that teaching quality audits inhibited innovation in approaches to teaching because they diverted academic staff time towards administrative tasks rather than seeking to bring about a change. It is also felt that conceptions of teaching quality vary across different disciplines/subjects - though this issue is not thoroughly explored in the literature.

3. International Models

3.1 Introduction

In a climate of increasing accountability in the public sector, governments in many countries have set up publicly funded quality assurance agencies to conduct external audits of higher education institutions (Lee, Huang, and Zhong 2012 ;Yin et al. 2014). While at the beginning of the 1990s, only approximately 50% of European countries had introduced quality assessment activities, by 2003, all except Greece had entered some form of national assessment (Schwarz and Westerheijden (2004). Furthermore, the Bologna Declaration (1999), and its emphasis on the need to develop comparable criteria and methodologies for assuring quality in higher education, also engendered new developments in the field, including the drafting of the Standards and Guidelines for Quality Assurance in the European Higher Education Area and the establishment of the European Quality Assurance Register for Higher Education (Cardoso et al 2015).

The aim of these audits generally is to: improve quality assurance in the core functions of the higher education institutions, including teaching and learning; improve the competitiveness and reputation of higher education; ensure that the institutions are accountable for the quality education; and ensure that public funding of institutions is transparent and contributing to the national and regional economy (Shah 2012). Burke (2005) defined four basic approaches to quality assurance: accreditation approaches, which were developed in the US and focus on institutional development and improvement through self-study and peer-review; audit approaches, which were developed in the UK and examine institutions' quality assurance and improvement processes rather than the delivered quality of the education itself; quality assessment or evaluation approaches, which examine the effect of an institution on its students by looking at learning outcomes; and external examination approaches, in which experts review the performance of individual students, of study programmes or of entire institutions.

In this section, we review the international literature on external quality audits of higher education institutions and then examine studies of quality assurance systems that are specifically designed to identify and measure of teaching quality across HE.

3.2 External audits of higher education institutions

While external audits have been in place for more than a decade in some countries, there is limited research on the extent to which such audits have improved quality assurance in universities (Shah 2012). Overall, there is a broad agreement between the studies that have been undertaken that external quality audits, together with internal university processes, have been a driver in improving universities' quality assurance processes. Studies by Saarinen (1995), Thune (1996), Smeby and Stensaker (1999), and Brennan and Shah (2000) indicate that self-assessment processes, as part of an external audit, can be successful in improving quality assurance processes in HE institutions.

A key function of external quality assurance is to stimulate change and improvement of teaching and learning (Cardoso et al 2015) and there is some limited evidence to suggest that this has happened in some countries. According to Shah (2012) research on external

quality audits in the New Zealand, Sweden and Hong Kong shows the effects have resulted in: increased responsibility for improving teaching and student learning; facilitating active discussion and cooperation within academic units on the means to improve teaching and student learning; clarification of responsibility for improving teaching and student learning in faculties; and providing information on the best practices both within the institution and across the systems (Shah 2012; also see Dill 2000). Wächter and Kelo (2015) concluded that quality assurance systems were there to guarantee compliance with minimum quality standards and support quality enhancement. Also, as quality assurance systems are generally run by not for profit organisations, their work can have legal implications for the HEIs they work with, even offering a licence to deliver.

A quantitative study undertaken by Stensaker et al (2010) in Norway suggests that national quality monitoring by the external agency had had a positive impact in several areas including: effects on the scholarly discussion on learning and teaching; effects on the quality of learning and teaching; and effects on staff engagement in learning and teaching questions (see also Gynnild 2007). Other studies in a range of countries have also found positive effects, including: enhanced learning and teaching, e.g. in the West Indies (Gift and Bell-Hutchinson 2007); the enhancement of learning and teaching, sequences of learning and the approach to student assessments (Gerbic and Kranenburg 2003; Horsburgh 1997); and in New Zealand, an improvement in teaching outcomes (Carr, Hamilton, and Mead 2005).

However, the research concerning the impact of external quality assurance schemes on teaching and learning is sparse and lacks depth (Cardoso et al (2015); Stensaker 2010; Newton 2002; Cartwright 2007). Hénard and Roseveare (2012) allude to the lack of quantitative indicators of quality teaching, but stress this should not stop attempts to assess and interpret the impact of good quality teaching. They look at this in some depth when considering one of their policy levers, assessing impact. At the time of publication, they considered that there was evidence that support for developing quality teaching was being assessed, but far less that evaluation of teaching was being undertaken, and if it was, it was unlikely to be used to improve quality.

Recently, Cardoso et al (2015) noted that "research in this area still seems to overlook the higher education institutions micro processes, which quality assurance is supposed to improve (see Kohoutek 2014). Consequently, it is difficult to trace the concrete effects external quality assurance has on higher education, including, the levers that stimulate improvements in teaching and learning. Cardoso et al (2015) conclude that "It seems that the institutional consequences of quality assurance have not yet contributed much to actual improvements in teaching and learning, or to transforming the student learning experience (see also Harvey and Newton 2004; Rosa, Tavares, and Amaral 2006). This is perhaps one of the reasons why we currently see considerable interest in integrating learning outcomes in existing external quality assurance schemes, where both evaluations and/or accreditations use learning outcomes as the key criteria for assessing quality (Kohoutek 2009, 2014; Stensaker 2014; Cardoso et al 2015).

3.3 Measuring teaching quality across HE

Below we consider studies which have systematically examined attempts to identify and measure teaching quality across HE in China, Australia and Germany. The research on the evaluation of teaching quality across the HE system in China is of especial interest

because it is on-going and sheds light on key issues that are relevant in relation to England and other countries.

3.3.1 China

On-going research on the Undergraduate Teaching Evaluation (UTE) programme in China highlights some potentially valuable lessons for England. In China, the quality assurance of undergraduate teaching is primarily represented by the UTE led by the Ministry of Education. The first five-year cycle (2003–2008) focused on teaching quality at the institutional level using evaluation criteria designed by the Ministry of Education. The criteria covered eight major indicators:

1) guiding principles of university operation;
2) teaching staff;
3) teaching conditions and the utilisation of teaching facilities;
4) subjects and teaching;
5) teaching management;
6) academic atmosphere;
7) learning outcomes; and
8) special features. (Yin and Wang, 2015:103)

All of the higher education institutions providing undergraduate programmes were evaluated using these criteria. The evaluation procedures were also standardised; they usually included self-evaluation, site visits by an expert panel and follow-up modifications. Based on the review written by the expert panel, the overall teaching quality of each institution was ranked as excellent, good, pass or unqualified (Yin and Wang, 2015).

During the first five-year cycle (2003–2008), 589 higher education institutions were evaluated by the Higher Education Evaluation Centre. According to the results released by the Centre, the numbers of institutions ranked as excellent, good and pass were 424 (72.0%), 144 (24.4%) and 21 (3.6%), respectively, and there were no unqualified institutions. According to Yin and Wang (2015:1033): "These results appeared inconsistent with the public perception of a decline in the quality of higher education, and the surprisingly high ratio of excellent institutions was widely questioned by scholars (Chen 2008; Jiang 2010)". Studies of the first round of UTE showed that it had led to both positive and negative changes. For example, Mohrman, Wang, and Li (2011) argued that, although the UTE provided an opportunity for institutions to seriously rethink their vision in the context of higher education expansion, and made them recognise the importance of undergraduate teaching, the complex criteria set by the Ministry of Education greatly limited the autonomy of higher education institutions. Furthermore, the competence of the on-site evaluators and the specialists in the expert panel were extensively questioned. Liu (2013) found that the effect of UTE on resource commitment and quality management was more significant than its effect on other dimensions, including teaching, learning and teaching-research balance.

More importantly, although the UTE programme was designed to assess the quality of undergraduate teaching, Lee, Huang, and Zhong (2012) pointed out that the information collected in the first cycle was mainly about the 'macro factors' influencing university teaching, such as facilities, equipment and the regulation of teaching administration rather than the 'micro elements' related to the real quality of teaching and learning, such as instructors' teaching strategies and the characteristics of university students' learning.

Moreover, although undergraduate students were supposed to be the beneficiaries of the UTE, they had little voice in the quality assurance system, and they appealed to increase their presence in the UTE process (Mohrman, Wang, and Li 2011). The first round of UTE has been extensively criticised for its neglect of student participation, the actuality of learning and teaching and the autonomy of higher education institutions (Lee, Huang, and Zhong 2012; Mohrman, Wang, and Li 2011). The perceived failure of the UTE to reveal the effect of the institutions on students, especially on their learning outcomes led to a strong current of opinion that the evaluation should be more focused on the actual teaching processes in higher education institutions, and the measures of quality evaluation should be modified to incorporate student learning (Yin and Wang, 2015).

In early 2012, the second round of UTE was initiated. In the on-going, second round of UTE, higher education institutions are required to establish a self evaluation system that assesses the conditions, processes and effects of undergraduate teaching. Currently several researchers are exploring undergraduate teaching in China from the perspective of student learning and considering the implications for the improvement of (UTE and) teaching and teaching quality evaluation in higher education in China, and the development of the self evaluation systems required for the second round of UTE (Yin and Wang, 2015). For example, using the Course Experience Questionnaire, Yin and Wang, (2015) examined the quality of undergraduate teaching by investigating the relationships between students' course experience, the learning outcomes demonstrated by the students and the learning environment. They conclude:

> "First, higher education institutions in China should monitor the quality of undergraduate teaching and its effect on students in a conventional manner using the CEQ as a measure of internal evaluation. Second, the results of this study, in line with other scholars' observations (e.g. Lee, Huang, and Zhong 2012; Liu 2013), indicated that, when evaluating the conditions of institutions, UTE in China should pay more attention to the accessibility of learning resources offered by the institutions and their appropriateness for students' learning needs, rather than only considering the provision of equipment, services and facilities, as the former are more powerful in improving students' course experience" (Yin and Wang, 2015:1046).

The research on the UTE in China highlights the importance of extending the focus of national teaching quality evaluation programmes beyond the "macro factors' influencing university teaching, such as facilities, equipment and the regulation of teaching administration to encompass the 'micro elements' related to the real quality of teaching and learning, such as instructors' teaching strategies and the characteristics of university students' learning (Yin and Wang, 2015:1034). The research also emphasizes the importance of including the student voice and considering students' learning outcomes.

Noting that there is evidence that students in different disciplines have differing expectations from their courses (e.g. Parpala, Lindblom-Ylänne, and Rytkönen 2011), Yin and Wang (2015) suggest that future research in China should explore the possible impact of students' disciplinary background on their course evaluation.

3.3.2 Australia

In August 2004, the Australian Learning and Teaching Council (ALTC) was established as a national focus for the enhancement of learning and teaching in higher education in Australia (ALTC, 2009). The criteria for excellence in university teaching employed by this

federally funded body for the purposes of national recognition and reward are now widely employed across the sector. A growing number of Australian universities apply these criteria to judge the effectiveness of the teaching of their staff in institutional teaching awards and to assess the teaching component of promotion applications. In effect, the ALTC criteria have become accepted as the proxy list of skills and practices of effective university teaching in Australian higher education. (Devlin and Samarawickrema 2010) The ALTC 's five key guiding criteria for determining excellence in university teaching for the purposes of recognition and reward are:

1. Approaches to teaching that influence, motivate and inspire students to learn;
2. Development of curricula and resources that reflect a command of the field;
3. Approaches to assessment and feedback that foster independent learning;
4. Respect and support for the development of students as individuals; and
5. Scholarly activities that have influenced and enhanced learning and teaching (ALTC, 2008).

A study conducted by Devlin and Samarawickrema (2010) compared these ALTC criteria against the SEEQ dimensions (Marsh & Roche, 1994) of effective teaching to determine the degree of alignment between the two (Devlin and Samarawickrema 2010). They conclude that while "the criteria adopted by the ALTC have been a useful proxy of effective teaching and have been successfully used to recognise and reward effective teaching and to guide enhancement efforts (…) owing to massive and ongoing change in higher education, it is time for a renewal of these criteria and, through them, the national collective understanding of effective teaching in Australia" (Devlin and Samarawickrema 2010: 122). The changes in higher education that Devlin and Samarawickrema (2010:
119) highlighted included:

- Increased student diversity due to the massification and the internationalisation of Australian higher education, which means that effective teaching must be able to manage and address such diversity. In order to engage all students, teachers must have an appropriate pedagogical response that accommodates a wider range of both learning styles and preferences and a wider range of language, cultural and educational backgrounds than has previously been the case (Devlin, 2007c; Higgins, Hartley & Skelton, 2002; Skelton, 2002).

- Students' increased participation in paid work in order to manage the cost burden of higher education, which has led to a decrease in student class attendance and in the time given by students to other study related activities (James, Bexley, Devlin, & Marginson, 2007). There has also been a concurrent increase in institutions offering flexibility in course delivery, assessment tasks, including the integration of paid work into formal learning experiences. Providing for flexible, 'anytime-anywhere' education is increasingly an expectation of the effective university teacher.

- Changing expectations of employers and accreditation and professional bodies that impact on teaching and learning at the university level. Demands for particular skills and attributes of graduates from employers are common and accreditation requirements sometimes provide significant curriculum and teaching challenges that are continuously changing.

- Effective teaching in higher education is also linked to technological changes.

> Advances in this area have had significant impact on teaching and learning in recent years (Hannan & Silver, 2000), which in turn have both assisted teachers to reach larger numbers of students as well as manage diversity and added to the complexity of the tertiary teaching and learning environment (Devlin, 2007a). Staff must continually learn new skills and familiarise themselves with new ways of interacting and communicating with students and be capable of teaching on- campus, off-campus and in blended environments employing technologies and pedagogies suited to the context and student cohort (Benson & Samarawickrema, 2009).

Devlin and Samarawickrema (2010) conclude that collective understandings of effective teaching need to be periodically reviewed and renewed to absorb the transformations that are occurring within universities and beyond them. However, they also warned that

> "Any revision to the existing ALTC criteria of effective teaching would need to be rigorous, have a sound methodological and evidence base and be carried out by an independent entity that will make public the process through which the criteria are derived. Without such thoroughness, the results of the renewal exercise might have questionable validity and, perhaps, limited acceptance. Any new criteria would need to be articulated in precise language and provide sub-categories so that they are explicit and clearly understood, as is the case for the current criteria. It is envisaged that such a renewal would form part of an ongoing endeavour to ensure that future developments, trends, understandings, government directions, stakeholder expectations and student needs are continually considered and incorporated into the collective understanding of effective teaching. The notion of effective teaching in higher education can then continue to have resonance and meaning within a changed and changing context."

Devlin and Samarawickrema's study suggests that context is critical and that it is subject to continuous and multiple changes imposed by forces from within and outside universities. As a result our collective understanding of competent, professional and effective teaching must continually evolve in order that it accurately reflects and continually responds to the contexts in which learning and teaching is undertaken. Devlin and Samarawickrema also call for an ongoing agenda that continuously investigates and articulates the meaning of effective teaching in a changed, and changing, context.

3.3.3 Germany

Bornmann et al (2006) review the use of multi-stage evaluation procedures as the main quality assurance instrument for evaluation of teaching and learning in HEIs in Germany. In the multi-stage procedure, academic review begins with internal self-assessment, whereby an academic programme or institute conducts its own analysis of strengths and weaknesses for a self-evaluation report. The next step is external evaluation. Here peer reviewers conduct a site visit of the programmes or units under evaluation and prepare an external evaluation report. The follow-up stage entails implementation of the reviewers' recommendation. Bornmann et al. (2006) undertook a comprehensive investigation of the acceptance and success of this evaluation procedure at German HEIs, focusing on the two most tried-and-tested and best-known evaluation procedures in Germany: VNU – a consortium of six German universities for evaluation of teaching and studies - and ZEvA – a common agency for Lower Saxony HEIs. The evaluation procedures used by ZEvA and VNU do not aim to produce a ranking of the participating institutions.

The study was based on a mail survey by questionnaire of all former external reviewers and members of institutes who participated in evaluations conducted by ZEvA and VNU Bornmann et al (2006) also conducted content analyses of the reviewers' recommendations in the total of 203 external reviewers' reports on completed evaluations of 28 HEIs and 25 disciplines produced by ZEvA and VNU up to the end of 2001.

The results of Bornmann et al's (2006) study show that the multi-stage procedure used to evaluate teaching and learning in HEIs in Germany enjoyed wide acceptance and were seen to be useful and effective. However, suggestions for improvement were made in relation to several areas. In particular, it was noted that the reviewers' recommendations were unevenly distributed across the various areas. A large part of the recommendations address the areas ''planning and organization of teaching and learning'' and ''resources'', while recommendations considering ''goals for teaching and learning'' were much more rare (Bornmann et al 2006). Concerns were also raised about the heavy work burden that the process entails, about the cost-benefit value of the evaluation process, and the composition of the panel of external reviewers for evaluating teaching and learning. With regard to the latter, It was suggested that the panel of reviewers should (be broadened) include an expert from another country, an expert working in professional practice, a student or graduate and possibly a representative of the non-professorial scientific staff. The inclusion of a student on the review panel was viewed as a means of increasing students' acceptance of the evaluation process. Bornmann et al (2006) conclude that against this background, it makes sense to perform regular quality assessments of the procedures for quality assurance and improvement", quoting (Vroeijenstijn 2000, p. 66):

> ''One thing is for sure, one never can say that a system for external quality assessment is finished. It always can be improved and it has always to be changed for keeping the academic world alert and to prevent quality assessment becoming a ritual dance'' (see also Vroeijenstijn 1995, p. 38; and European Training Foundation 2000, p. 24).

3.4 Summary

Governments in many countries have set up quality assurance agencies to conduct external audits (including assessment of teaching and learning) of higher education institutions. Currently, there is limited research on the extent to which such audits have improved quality assurance in universities. However, studies suggest that external quality audits, together with internal university processes, have stimulated change and improvement of teaching and learning in some countries. Research shows the importance of considering the 'micro processes' surrounding teaching and learning, such as lecturers' teaching strategies and the characteristics of university students' learning, however, there is very little research on the impact of audits by external agencies on these. There is an emerging interest in how students in different disciplines have differing expectations from their courses but it remains unclear whether and how this impacts on their perceptions of teaching quality and course evaluations.

Research on quality assurance systems designed to measure teaching quality across HE in China and Germany points to the importance of a three-stage process involving self- evaluation by HEIs (stage one), external evaluation by peer reviewers (stage two), and implementation of peer reviewers' recommendations (stage 3). In China, the ranking of the overall teaching quality of each institution as excellent, good, pass or unqualified during

the first five-year cycle of evaluations of teaching quality across HE (2003-2008) was widely challenged by scholars due the high number of HEIs that were ranked as excellent. The first cycle of evaluations of teaching quality was also criticized for failing to take students' learning outcomes into account. As a result, the second cycle of evaluations (2012-ongoing) is placing a greater emphasis on actual teaching processes in HEIs, and the measures of quality are being modified to incorporate student learning.

The study of systems used to evaluate teaching quality in HEIs in Germany reveals that they enjoyed wide acceptance and were seen to be useful and effective. However concerns were raised about the need to: reduce the work burden associated with the systems; broaden the composition of expert panels (to include, for example, students, international experts, and people working in professional practice); and ensure that reviewers' recommendations were more evenly distributed between 'planning and organization of teaching and learning' and 'resources' (which were common) and 'goals for teaching and learning'' (which were rare).

Research on the Australian Learning and Teaching Council's criteria for excellence in university teaching indicates that criteria and metrics used to identify and measure effective teaching need to be continually reviewed in order to ensure that they reflect changing definitions of effective teaching due to changing students needs, new modes of delivery, changing employer needs and technological innovation.

4. Teaching excellence in HE

4.1 Introduction

There is an extensive literature on teaching excellence in HE (Gunn and Fiske 2013; Fanghanel et al 2015). This frequently concerns teaching portfolios or personal claims of excellence", and is often related to monetary reward or promotion (Bradley et al 2015). Many of these schemes fall under the category of teaching excellence and individual claims or nominations of excellence, such as the distinguished teachers awards in the US (Elton 1998), and National Teaching Fellowships in the UK (Skelton 2007). In this section, we provide an overview of the key themes in UK and international literature on teaching excellence in HE and then examine issues surrounding the measurement of teaching excellence..

4.2 Defining excellence

There is a striking lack of consensus in the literature on how to define teaching excellence in higher education (Gunn and Fisk 2013; Land and Gordon 2015; Little et al 2007; Skelton 2007, 2009). It is a contested concept (Land and Gordon 2015; Macfarlane 2007; Skelton 2009), with varied competing interpretations (Madriaga and Morley, 2016) and terminology (Gunn and Fisk 2013; Tsui 2013), and it is also regarded by some as being potentially divisive (Gibbs 2007). Grifoll (2014) comments that definitions of excellence in higher education may well depend on the person using the term and their motivation for doing so. In a study of the changing role and marketing practices of universities in a global and knowledge driven economy Gaspard (2013) found that the concept of excellence is widely used on university websites without definition to promote institutions or aspects of their offer.

In a recent report for the European Association for Quality Assurance in Higher Education (ENQA), Brusoni et al. (2014) examine several aspects of excellence, including in management, research, teaching and student performance. Their review of excellence in teaching revolves around many of the arguments that have recently been well rehearsed by the likes of Gibbs (2010, 2016) and Gunn and Fiske (2013), and there are a number of important issues concerning the definition of teaching excellence that arise. Brusoni et al (2014: 12) note that:

> "Excellent teaching is determined by factors such as the inspirational nature of individual lecturers, the organisation of presentations, the interaction with students as participants and how well the information provided meets the learning objectives of the course. Excellence can be identified both in terms of student satisfaction and also in terms of the performance of students in assessment. There are differences between deep and surface learning. Excellent teaching may be seen as the efficient presentation of information which maximises the students' opportunities to gain the highest marks from the course. Alternatively, excellence could be recognised as the stimulus for students to engage with the subject and to enhance their understanding and knowledge".

As Brusoni et al (2014: 13) observe, the situation is further complicated by the fact that excellent teaching is not only down to the individual teachers and the learners they work with, but also the supporting resources and systems, which facilitate the learning process (see also, Elton 1998; Gibbs 2012; Gunn and Fiske 2013). Here Brusoni et al (2014: 13)
quote Elton (1998):

> "Individual teaching excellence is a necessary, but not a sufficient condition for an excellent student learning experience. In addition there must be excellences at departmental and institutional levels. They can however be developed on the foundation of individual excellence".

Similarly, Gunn and Fiske (2013: 19) differentiate between conceptions of excellence related to individual teachers' philosophies and practices and conceptions of excellence related to sector, institutions, and disciplines. A summary of the distinction between individual and system-based conceptions of excellence is provided in Figure 1.

Figure 1: Individual and system-based conceptions of excellence

Individual teacher excellence	Conceptions of excellence related to the practices of individual teachers, which focuses on lecturing and in some cases other activities such as the provision of feedback to students.
System-wide conceptions of excellence	Conceptions of excellence related to the systems and resources which facilitate the activities of individual teachers, which may be focused at departmental, institutional, disciplinary and sector levels.

Excellence can also be interpreted in either 'competitive' and 'elitist' or 'egalitarian' and 'inclusive' terms. Strike (1985) distinguishes between norm-referenced and criterion- referenced conceptions of excellence. Norm-referenced conceptions of excellence define excellence relative to the performance of others and imply competition and elitism because " people will be in competition for excellence and, as a matter of logic, not everyone can attain it". In contrast, criterion-referenced definitions of excellence "define excellence in relation to a standard such that people are not in competition for it and, in principle, if not necessarily in fact, everyone can attain it".

Figure 2: Norm-referenced and criterion-referenced conceptions of excellence

Norm-referenced conceptions of excellence	Excellence is defined relative to the performance of others
Criterion-referenced conceptions of excellence	Excellence is defined in terms of a standard which in principle can be attained by all

Norm-referenced conceptions of excellence	Excellence is defined relative to the performance of others
	concerned.

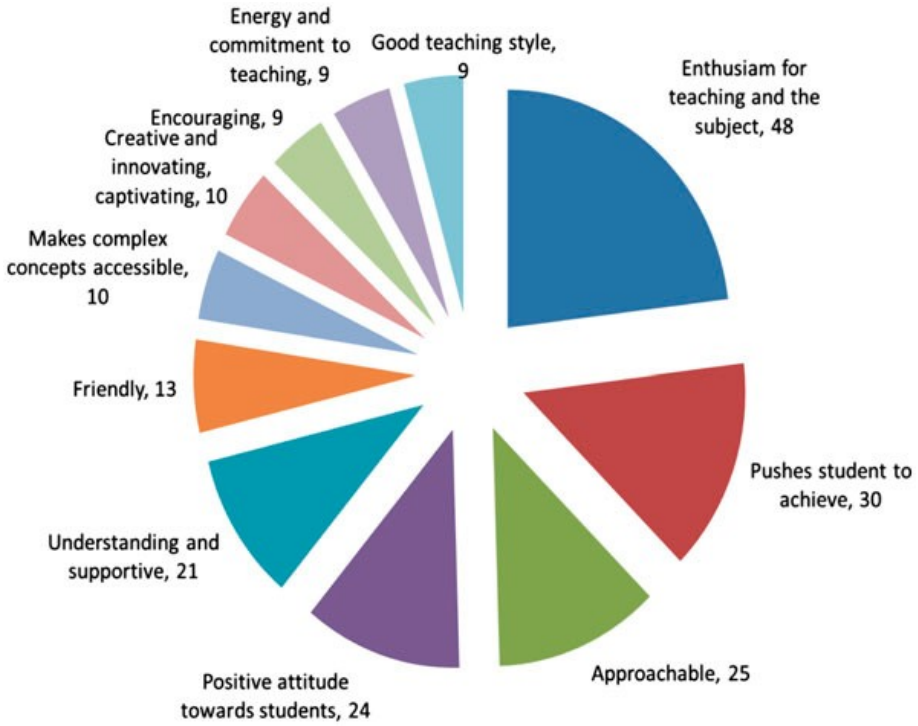
Strike favours a criterion-referenced approach on the grounds that it is more equitable and inclusive. This view chimes with the notion that definitions of teaching excellence (and teaching quality) in higher education should incorporate different modalities according to the purposes, missions and values of different HEIs (Little and Locke 2011).

4.3 Perceptions of excellence

There are few empirical studies on academics' views about teaching excellence and what they regard as good indicators of excellent teaching (Cardoso et al 2015; De Courcy 2015), perhaps because of the difficulties associated with defining the term precisely (Harvey and Green 1993). Moreover, although there has been a growth in student- nominated teaching excellence awards, there is relatively little research on student perceptions of teaching excellence (Bradley et al 2015; see also: Carusetta 2001; Hill, Lomas, & MacGregor 2003). Recently, however, a handful of studies have emerged which look systematically at teaching excellence as perceived by students. These concentrate on individual teacher excellence and, in so doing, largely focus on lecturing and give little or no consideration to other elements of teaching, such as feedback.

In addressing the gap in evidence in relation to students' perceptions of teaching excellence (whilst at the same time presenting some of the challenges in managing a student-nominated teaching awards scheme), Bradley et al (2015) offer some rich insights into this issue. Their study involves an analysis of student comments derived from an internal survey for a student-nominated inspirational teaching awards scheme at a large university in the United Kingdom (UK). Figure 3 below shows the qualities that students identified as associated with inspirational teaching. These reveal that students want to be taught by staff who are enthusiastic about their subject area, academically stimulating with up-to-date knowledge of their field, but who are also empathetic and wish their students to develop their full potential. (Bradeley et al (2015: 239).

Figure 2: Student perceptions of teaching excellence



Source: Bradeley et al (2015)

Bradley et al (2015) note that the qualities identified by the students in their study map on to many of the qualities of good teaching expressed by Ramsden, Margetson, Martin and Clarke (1995 and cited by Trigwell (2001):

- Enthusiasm for subject and desire to share with students;
- Modifying teaching to suit the students, subject matter and environment;
- Encourage learning for understanding and concern about the development of students' critical thinking;
- Ability to transform and extend knowledge;
- Show respect for their students, interest in professional and personal growth with sustained high expectations.
  (Bradeley et al 2015: 238}

The qualities identified by the students in Bradely et al's (2014) study are also consistent with earlier studies by Brown & Atkins (1993), Hill et al. (2003) and Lowman & Mathie (1993). Interestingly, they are also consistent with the findings of Sherman et al (1987)'s review of the literature on teaching excellence in the late 1980s which identifies five characteristics that were regularly and consistently attributed to teachers classified as excellent: enthusiasm, clarity, preparation and organisation, stimulating, and love of knowledge (cited in Brusoni et al 2014). However, while there are considerable overlaps between the qualities associated with individual teacher excellence, these qualities are generally ill-defined, difficult to describe and difficult to assess (Brusoni 2014:25-26).

Other research into student-led teaching awards suggests that student perceptions of teaching excellence are divided into two main areas: the personal attributes and style of the lecturer or tutor, and the content, structure and delivery of sessions (Davies, Hope & Robertson 2012). Gunn and Fiske (2013:36) note that alumni also provide a useful retrospective reflection on what constituted excellent teaching (Hammer et al 2010). For example, a qualitative analysis of feedback from one group of alumni at the University of Limerick, illustrated that approachability, interest in students, helpfulness and patience were the most recollected traits of academics considered as excellent teachers (Moore & Kuol 2007). Although a command over disciplinary areas was considered important by this group, it was viewed more as an outcome of an effective teacher-learner relationship than as the central ingredient in excellence (Moore & Kuol 2007: 140).

4.3 Measuring Excellence

ENQA's report on the Concept of Excellence in Higher Education reports that Quality Assurance systems for Higher Education across the world are not particularly geared to identifying or promoting excellence, but are more about evaluating quality and recommending areas for improvement. They set thresholds for quality and encourage improvement. There are some exceptions for Centres for Excellence, but these are related to research and SoTL activity. Their view is that if Quality Assurance organisations were to become involved in identifying and certifying excellence, there would need to be very precise definitions and clear distinctions drawn between excellence in teaching and excellence in research Brusoni et al (2014).

Student-led teaching award schemes is an area in which there is much debate around the issue of measuring teaching excellence. As noted above, students rating and nominating their lecturers for university teaching awards fall under the category of teaching excellence (Bradley et al.2015). For student-led teaching awards, the dominant methods to judge teaching excellence have been teaching observation and evaluation questionnaires, which has raised questions about whether they are a popularity contest and whether students are capable to identify what constitutes good teaching practice (Aleamoni 1999; Davies, Hope, and Robertson 2012; Gibbs 2012; Macfarlane 2007). Despite these criticisms there remains a steady effort to operationalize notions of teaching excellence, especially in the form of student-led teaching awards schemes (Madriaga and Morley 2016). Bradeley et al. (2015: 240) suggest that while other measures, such as the National Student Survey results, are more likely to be used to judge teaching quality by an institution than a local award scheme, "if student comments in the internal survey are deemed to be an unreliable measure on which to base student-nominated teaching awards, then questions of reliability about student comments in module evaluations or the National Student Survey can be made."

In their review of the literature on teaching excellence, Gunn and Fisk (2013: 20) highlight a tension between competing discourses within the 'teaching excellence" literature as revealed in the findings here, particularly a discourse of cynicism where teaching excellence is a facet of neo-liberalism as opposed to a discourse of pragmatism in which institutions demonstrate teaching excellence to convince stakeholders (Ball 2003; McGettigan 2013; Madriaga and Morley, 2016: 172). However, in their study of a UK university, Madriaga and Morley (2016) found both of these discourses are interlinked. They found that while there has been much criticism leveled at the institution's teaching award scheme, many academics support the need for the institution to recognise and

value teaching within the institution:

> "It is not the rationale of the scheme that is contentious. It is the current method in place at the institution" (Madriaga and Morley 2016: 172).

4.4 Summary

ENQA's report on the Concept of Excellence in Higher Education reports that Quality Assurance systems for Higher Education across the world are not particularly geared to identifying or promoting excellence, but are more about setting thresholds for quality and encouraging improvement.

Defining excellence is considered to be a challenge and the literature provides a plethora of definitions of excellence. It is a contested concept, with varied competing interpretations**.** The situation is complicated by the fact that excellent teaching is not only down to the individual teachers and the learners they work with, but also the supporting resources and systems, which facilitate the learning process. Excellence can also be interpreted using either norm-referenced definitions of excellence, which define excellence relative to the performance of others and imply competition and elitism, or criterion- referenced definitions of excellence, which define excellence in relation to a standard such that people are not in competition for it.

There are few empirical studies on academics' views about teaching excellence and what they regard as good indicators of excellent teaching, perhaps because of the difficulties associated with defining the term precisely. However, studies of teaching excellence as perceived by students indicate that students want to be taught by staff who are enthusiastic and knowledgeable about their subject, empathetic, approachable, helpful and patient, and encourage students to develop their full potential.

5. Metrics

5.1 Introduction

There are many potential metrics for measuring teaching quality in HE. Berk (2005) offers a list of 12 sources of evidence of teaching effectiveness: Student ratings; peer ratings; self-evaluation; videos; student interviews; alumni ratings; employer ratings; administrator ratings; teaching scholarship; teaching awards; learning outcome measures; and teaching portfolio. Many of these will be essential for individual teachers to review and contemplate when reviewing their own performance; for managing teams within an HEI and encouraging enhancement, they will be invaluable. For easily comparable metrics across institutions however, they are less valuable, as self-evaluation, videos, student interviews, teaching scholarship, and teaching portfolio are likely to be qualitative and institution specific. While student ratings, peer ratings, alumni ratings, and administrator ratings and likely to be defined by the institution and not measured on common scales. Teaching awards are often particular to the institution and accolades given to individual teachers for individual courses or programmes (see for example http://www.ethicsguidebook.ac.uk/Adult-Social-Services-staff-service-users-or-their-families-125).

Chalmers (2008) describes the extensive work conducted in Australia to develop a nationally agreed and valid series of Teaching Quality Indicators. This is a thorough approach based on: research into methods used across the globe; the meaning, purpose and nature of performance indicators; and the difficulties and challenges particular to qualitative and quantitative indicators. The barrage of indicators they use across a number of institutions work as a logic chain approach to understand inputs, processes, outputs and outcomes. Indicators tested included institution climate aspects such as adoption of student-centred learning perspective, diversity aspects such as active recruitment and admissions, assessment aspects such as commitment to formative assessment and engagement aspects such as fostering and facilitating learning communities. This approach is undoubtedly of value to individual institutions in improving the quality of teaching and learning, and enabling dialogue between peer institutions, assuming a number sign up to it. It is also in the context of a smaller sector of institutions than is found in England (43 HEIs compared to over 160 in the United Kingdom).

Both Berk (2005) and Chalmers (2008) identify a large number of potential metrics, but the assessment of teaching quality across HE requires a crisper approach, with a small number of metrics that are easily replicated across the sector and understood by a variety of stakeholders, as well as being valuable in driving improvements in teaching quality in the sector. In this section we look at quantitative metrics, including input and output measures, process measures and rankings. We also look at the arguments for qualitative measures, before looking at gaps in the evidence base. A table summarising the strengths and weaknesses of the metrics discussed in the literature is provided in Annex 1.

5.2 Quantitative metrics

5.2.1 Input and output metrics

In the United States attempts to measure teaching quality in higher education date back the 1960s. Massey (2010) reviews a number of metrics that could be used including price and outputs. He takes an economist's perspective on higher education and looks at measures of efficiency and productivity and the metrics to describe them in a higher education context. He discusses the challenges of developing input and output measures that capture the required complexity and weightings to avoid distortion and gaming. He feels that the developers of metrics need to build in a quality element.

Massey (2010) rejects both price (an input measure) and outputs as suitable stand alone metrics. Pricing is rejected due to the distortions of high fees and highly selective private universities at one end of the scale and state subsidies and tuition ceilings in public universities at the other in the United States. For outputs, a number of output-quality systems exist in the United States, including the Collegiate Learning Assessment, the National Survey of Student Engagement (NSSE), the Educational Testing Service (ETS) Proficiency Profile and ACT's Collegiate Assessment of Academic Proficiency (CAAP) (Massey 2010). However, Massey observes that while these have been accepted as part of the voluntary system of accountability, no single metric has provided consensus on being reliable and comparable for whole institutions, in part due to the highly commercialised HE sector in the country (Massey 2010).

Fearing that tightening budgets and unhelpful metrics would allow or even encourage universities to do less with less, Massey argues that productivity and efficiency metrics need to be supplemented with a robust, external review of quality. Massey prefers some form of academic audit, citing the UK system, but ensuring freedom from political interference and allowing sufficient time for the method to be embedded in the American HE system.

Another potential metric system is offered by the French website www.letudiant.fr which holds listings of learning opportunities including sixth form (lycée) and higher education. This is said to cover 5,073 training organisations, 42,894 training opportunities and 33,340 evaluations, and allows prospective students to find courses that are of interest and see how they are regarded by those who have undertaken them. Ratings can be given by teaching staff, students or parents, and the ratings style is comparable to Trip Advisor, with institutions able to respond to ratings.

The data held on the website is considered by l'Etudiant to be a key source of information for young people as they prepare their requests for admission to University and to particular courses. More officially, the ratings are given little credence due to the small numbers of students providing ratings even for programmes with large numbers of students. Some of the measures might be considered "hygiene factors" in terms of key resources such as disabled access and library facilities, while the nearest aspect of the above to assessing the quality of teaching is Equipe éducative.

Gibbs (2010) advocates class size as one of a basket of metrics for teaching quality. Cuseo's (undated) extensive review into class size as a factor in learning outcomes, finds that large class sizes mean more reliance on a lecture style delivery, reducing the student

engagement with the subject, constraining the ambition of the stated learning aims of the course, limiting student satisfaction and learning outcomes. The study is mainly concerned with America. At whole institution level, it would be challenging to develop a metric for how much delivery is via large group lecture, and this is likely to vary in any case by programme, making the measure relatively meaningless unless that level of granularity was built in. However Soilemetzidis et al (2014) found that class size is important in student engagement and satisfaction via the UK Student Academic Experience Survey, so further exploration of how this could be made into a useful metric would be of value.

Another metric identified by Soilemetzidis et al is contact hours, with greater student satisfaction linked to a larger number of contact hours. While there is scope for using this as a metric, it has to be linked to programme of study, as humanities subjects are likely to have many more hours of self-guided study, whereas subjects with laboratory time for example will have more student contact hours. The Student Academic Experience Survey has not been analysed at institution level, as the sample was not drawn to be representative in that way, and there are not enough responses by institution to allow a robust analysis.

With the needs of Governing Boards in mind, Chaffee (2014) describes a project with 8 Colleges and Universities funded by the Teagle Foundation looking at indicators of educational quality. He found that standardised tests are difficult to implement, creating authentic assessments and metrics is expensive, students and institutions are diverse, so that it may need an institution by institution response. He describes educational outcomes as being defined as content knowledge, writing ability and critical thinking proficiencies. Assessment of these used direct evidence from their work during learning, as well as indirect evidence from students, alumni and employers. He acknowledges that learning is also influenced by other factors including aspects of inputs. He also argues that accreditation of the institution could be an indicator but generally reaccreditation activities are too infrequent. As an externally validated output measure, he contends professional exam pass rates might have some potential. Overall, he recommends a Dashboard approach with a range of indicators as shown in Figure 3 below.

Figure 4: Dashboard approach

Sample Board Indicators of Educational Quality

Inputs	Educational Process	Student Learning Outcomes
Student characteristics	Retention and graduation rates	**Direct Measures:**
Faculty characteristics	Student/faculty ratio	Professional examinations pass rate
	Student satisfaction	Multiple measures of cultures, critical thinking, communication, other learning outcomes.
	Teaching effectiveness	
	% budget to instruction, academic support	Satisfactory/exemplary student outcomes in Gen Ed & major
	Academic program review	
		Indirect Measures:
		Graduate satisfaction, employer satisfaction
		Graduate placement rate
		Academic program improvements

Source: http://agb.org/trusteeship/2014/1/learning-metrics-how-can-we-know-students- know-what-they-are-supposed-know

In their response to the Higher Education Green Paper, Higher Education Policy Institute (HEPI) (2016) cite findings from a student survey in 2015 which shows that Quality Assurance Agency for Higher education (QAA) ratings are the least important in a range of characteristics considered by students when choosing a Higher Education establishment. By contrast, degree outcome (an output measure) was considered the most important factor (40% of students said this was important). For many commentators, this has actually been viewed positively, as with institutions setting their own standards and degree programmes, comparability of degree outcome between institution is challenging at best, and with alternative measures such as Grade Point Average (GPA) and Higher Education Achievement Reports (HEAR) now available, it is becoming even less clear.

Gibbs (2016) is not supportive of the two key indicators advocated by the HEPI survey – degree outcome and student employment post degree – as he feels both are more reflective of the student rather than university teaching quality. He considers both to be inadequate output indicators. Instead, Gibbs reflects on what would be bad metrics (outcomes, retention rates, employability) and good metrics (student engagement (although still in development), learning gain (but will take time) metrics for a TEF. Gibbs (2016) also notes that research-based reputation has driven estimates of teaching quality in the past (the purported 'elitist' bias in previous teaching quality regimes in the UK is discussed in section 2. In addition he points out that student satisfaction is not the same as learning gain, and what students want may not be what makes for a good learning experience.

5.2. 2 Process metrics

The process oriented measures favoured by Gibbs (2016) are challenging to define and to measure. Two of the key ones are learning gain and student engagement (as opposed to student satisfaction). At the international level, learning gain is being tested through the Organisation for Economic Co-operation and Development's (OCED) Assessment of Higher Education Learning Outcomes (AHELO), although England is not part of the group of countries working on that OECD project. AHELO aims to evaluate HE student performan ss diverse

OECD's AHELO is a reasoning and problem-based scenario assessment similar to the American Collegiate Learning Assessment. One criticism of AHELO is it is impossible to create a language based complex test that is comparable across a number of different languages. Other aspects that are regarded unfavourably regarding AHELO are that institutions with highly selective entrance criteria or ones that are towards the top of the research rankings might not actually do so well on learning gain measures.

Other criticisms of AHELO (Ashwin, 2016, unpublished) are that it is independent of discipline studied, but actually measuring learning gain really ought to take account of the requirements of the degree programmes studied. In work he has done with Sociology students, tracking their engagement with their subject, Ashwin (2014) found it was possible to understand learning gain by how the students' perception and engagement with the subject changed over time. However this would need much more extensive work to develop into any form of easily reproduced and metric based system.

Schleicher (2016) looked at the challenges of measuring added value for higher education students, taking account of the context of the Green Paper from BIS in his HEPI address. His view was that internationally there is a need to develop a system for measuring learning gain in higher education. In part this is because people want to understand the quality of universities, rather than assuming universities are quality. It also stems from the increased access to higher education and that in many countries access to HE also requires payment of fees. The globalisation of the higher education market is another factor that drives the need for metrics. While he gives no definitive measures, the principles are explored, along with an exhortation to find a way to make measuring learning gain happen, as the data and technologies to enable it become increasingly available and he hinted at the potential offered by learning analytics.

If measurement of learning gain is very much at the exploratory and pilot stage in England currently, what is the situation with student engagement? Mandernach (2015) defines student engagement, noting that learning begins with student engagement and that being entertained is not being engaged, but thinking is. There are three aspects to engagement by a student – cognitive, behavioural and affective. Because the paper took an essentially American look at the issue of student engagement, the main methods for measuring engagement at institution level include the National Survey of Student Engagement (NSSE). There are a number of related and spin-off survey instruments that exist, depending on the needs of the institution or faculty and the whole is not critically assessed, however Massey (2010) has looked at how well NSSE fits into a range of metrics used in the USA.

In considering student engagement, it needs to be clear that it is not about governance but about teaching and learning. Kandiko-Howson (2016) has done research into student engagement in the UK. Talking to students, she found they understood that they had to engage in learning and also did not want to be thought of as customers. They expected a number of hygiene factors in their experience such as booked lecture theatres and libraries as discussed above under input and output measures, but provided these exist, they will not noticeably affect the quality of the student experience. Students do also accept, in her experience, that past student achievement is not indicative of current or future student achievement, so that they are able to contextualise high or low post- graduation employment rates.

Kandiko-Howson examined the UK Engagement Survey, developed and run by the Higher Education Academy (HEA) and now in its third year with slowly increasing number of institutions taking up the opportunity. Some qualitative testing of the questions showed that students found it a better feedback experience than the National Student Survey (NSS). The NSS has been reviewed and may in future have some student engagement questions, but she feels it is clear from the past that the NSS is subject to gaming both by institutions and by students who complete it, and so the changes may not help. She also discussed learning gain and the projects that HEFCE and BIS are undertaking now to look at that, but which will not be able to fit into the TEF for a number of years.

Buckley (2014) reported on the 2014 iteration of the UK Engagement Survey, including the challenges of adapting the NSSE questionnaire to a British environment. However given the perception among respondents that the survey is more relevant to them and their experience than the NSS, and that it attempts to assess the important issue of student engagement, its continued development is of major interest. This is particularly underlined

by the finding in the UK Student Academic Experience Survey (Soilemetzidis et al. 2014) that where students were not getting the experience they expected, it was in part due to their own lack of engagement with learning. However McCormick (in Grove 2015) has cautioned against adapting a tool that he oversees in the United States to drive internal institutional improvements, to be the basis of a national ranking.

Views on student evaluations and their contribution to teaching excellence metrics vary. Barre (2015) conducted detailed research into student evaluations and noted that there were elements that introduced bias into evaluations, and when known, can be adjusted for. These include student motivation, student effort, class size and discipline, but do not include gender, grades and workload. Barre concluded that while student evaluations might include some bias and might justify some criticism, they were also the best tool currently available for understanding teaching quality in higher education.

Endrizzi (2014), in her review of the literature from the United States and Europe on teaching quality, looks at using student feedback to understand quality. She feels that the theory of asking students about their learning experience is not problematic, however she has concerns that it could turn into a bureaucratic exercise with no dialogue, and that it can ignore the co-production nature of higher education, where learning is a joint process with responsibilities on the student as well as the teacher. Her belief is that at present the student body is taking the process seriously through the EU project QUEST. She highlights the American NSSE and the Australian Course Experience Questionnaire (CEQ) as good examples of student questionnaires. She also notes the NSS has been developed on the same lines as the CEQ, but also notes Gibbs (2010) has issues with aspects of this.

While student surveys are considered to be a reasonable and relatively serious way of assessing teaching quality, Endrizzi (2014) notes that there are some proven circumstances where poor ratings will be given, but this is to be expected (the example is given of a new teacher, with a large class, taking a generic subject which is compulsory, and where the learners are in the early stage of their higher education career). However where other tutor characteristics can be reviewed, gender seems to have no impact on the rating given, while good communication skills or being a warm and sympathetic person do not automatically also mean that the quality of the teaching is rated highly. Giving lenient marks is also not necessarily rated highly. However there remains some scepticism that it is still possible to influence ratings, and the degree to which this is possible is linked to whether authors are in favour or against student evaluation ratings. The finding on gender and bias in survey results has been challenged by Holroyd and Saul (2016) who have found evidence of consistently lower scores for women staff in the NSS.

Overall, Endrizzi (2014) notes that student perspectives are a useful way of understanding teaching quality, but that consideration has to be given to why the evaluation is done, by whom (the teacher themselves or the administration of the institution or even an external agency), at what stage in the learning process, using what questions, how it is administered and what the stated purpose of the exercise actually is.

Yorke et al (2013) raise the challenge of one set of metrics for all disciplines, in their work which highlights potential bias in the National Student Survey against Fine Art and Design courses. This would have a negative impact on specialist institutions in any Teaching Excellence Framework unless clearly understood by those using the data.

Within the field of student engagement research, there has been some consideration of learning analytics. In an interview in July 2015, JISC's Phil Richards (2015) looked at how learning analytics might be able to help prove teaching excellence. The existing applications which are discussed are not, however, metrics of teaching excellence but rather projects that can help teaching staff to engage more effectively with students and to improve outcomes for individuals. So proof of using learning analytics in a constructive and ethical manner might be a facet of teaching excellence but not a metric or a source of metrics in itself. This use of analytics at the local and programme level to support teaching staff is regarded by Ashwin as the way in which analytics can support teaching excellence as well, but it is not seen as a route for measuring teaching excellence.

Some care needs to be taken when considering student engagement as is clear from a Universities' UK blog[1] which links student engagement with projects looking at learning analytics. Given a wide definition of engagement, while analytics might be able to offer some insight at the individual level on behavioural and possibly cognitive engagement, affective engagement and the extent to which the student internalises the subject in their way of thinking and looking at the world, will not be touched upon.

5.2.3 Rankings

One way that excellence has been attributed to universities has been through the use of various ranking systems and the standing of individual institutions can be cited in promotional materials and on their websites, including whether they are in the top 10, top 100, top 200 or other banding. The international rankings have also driven higher education policy in some countries (Pol, 2012), so the use of ranking as a metric is worth consideration.

The Times Higher Education World University rankings[2] judge research intensive universities on all their main functions including teaching. Performance indicators are grouped into five areas: teaching; research; citations; international outlook; and industrial income. The methodology evolves over time as the producers aim to address criticisms and challenges to the ranking. The teaching domain provides 30% of the ranking and is based on the following indicators: reputation survey, staff to student ratio, doctorate to bachelors' ratio, doctorates awarded to academic staff ratio, and institutional income. Institutions provide and sign off the data for use in the rankings calculations, and data is manipulated to come up with the final indicators.

The Shanghai Academic Ranking of World Universities (ARWU)[3] has been produced since 2003 and ranks 1,200 institutions per year, although only the top 500 are published. The system "uses six objective indicators to rank world universities, including the number of alumni and staff winning Nobel Prizes and Fields Medals, number of highly cited researchers selected by Thomson Reuters, number of articles published in journals of Nature and Science, number of articles indexed in Science Citation Index - Expanded and

---

[1] http://blog.universitiesuk.ac.uk/2015/09/09/teaching-excellence-in-universities-can-it-be-defined-and- measured/

[2] https://www.timeshighereducation.com/news/ranking-methodology-2016

[3] http://www.shanghairanking.com/aboutarwu.html

Social Sciences Citation Index, and per capita performance of a university." This is very firmly based on research performance therefore, with little direct information on teaching.

The Quacquarelli Symonds (QS) World University Rankings[4] include a number of indicators: academic reputation from their global survey, faculty/student ratio, citations per faculty from Scopus, employer reputation from their global survey, proportion of international students and proportion of international faculty. Each indicator is subject to scrutiny and analysis before the final rankings are developed. The rankings are discussed in some detail by a number of authors.

Bouchard (2013) undertook a review of ranking systems for higher education and research. She found that ranking systems served more as a regulatory tool rather than a mirror to reality, or if considered as a mirror, rather as a distorting mirror. She concluded that rankings did not just represent social reality; they created the social reality to which they belonged and modified the behaviour of actors within it. In this respect, she pre- empted some of the findings of Wächter and Kelo (2015).

Yelland and Vale (2013) highlight four ranking systems – the Shanghai Academic ranking of World Universities, Times Higher Education (THE) World University Rankings, QS World University Rankings and U-Multirank – that have been developed to help individuals and organisations make decisions about higher education. However they consider that these rankings do little to offer information on the quality of teaching at any institution.

Altbach (2015b) discussed the use of rankings as a way of determining excellence at the international level among higher education institutions. He set out the history of the international rankings and the fact that the Shanghai and Times Higher Education rankings have both achieved a certain level of credibility at international level. He argued that rankings are inevitable, even for higher education, and that they are regularly used by Governments and "customers" to determine who is best and what is likely to be a better investment. Also for institutions themselves it is a good way to market themselves and improve their likelihood of "winning" the best students, who have an international marketplace to choose from.

His challenge to rankings was not in principle but in practice, as some part of the ranking process was based not on clear criteria but on some aspects that could be considered flawed. Thus rankings can be covertly skewed by the mix of science to humanities subjects (there is more pressure and potential to publish for the sciences), by the country hosting the institution and their native language(s) (the most prestigious publications publish in English and will favour papers authored in that language), and at the international level are generally driven by research prestige as there are no internationally accepted measures of teaching excellence.

Altbach also felt that there is a self-fulfilling element to the high rankings of the older and English speaking establishments. He therefore saw that while rankings exist and will continue to exist, those producing them have a challenge to ensure that the rankings are

---

[4] http://www.iu.qs.com/university-rankings/world-university-rankings/

based on clearly defined and understood measures, appropriate to what is being termed as excellence.

Westerheijden (2015) presented the U-Multirank system at las Palmas de Gran Canaria in November 2015, describing the perceived advantages of this multi-faceted ranking mechanism, depending on the needs of the user. He maintained that no ranking can meet the needs of all users, and offered a mechanism that can be adapted to meet the needs of the user. This is an international system, drawing data from over 1,000 institutions world- wide, and allowing users, including prospective students to use the tool to aid decision making.

The system is still in development and has received European funding for four years, but is intended to be self-sustaining in future, and independent. The system does in theory address the challenges to rankings that regard them as biased towards the old and traditional universities of Anglophone countries, but may not contain information on enough British Universities currently to be considered valuable for home students.

5.3 Qualitative metrics

While there is extensive research available on quantitative metrics of teaching excellence, there is less available on qualitative metrics although there are extensive discussions on the nature of quality systems and quality judgements. Thomas et al (2014) reviewed 27 papers on peer review of teachers and teaching in Higher Education, considering its potential as a way to improve quality of teaching. One conclusion arising from the paper is that peer review is a good tool for improving teaching quality; however it is also challenging to see how this might be turned into a metric for demonstrating excellence externally. Thomas et al. (2014) do regard this as an area for consideration however, as it overcomes some of the perceived bias that can arise from student evaluation of teaching. Massey (2010) did advocate that any quantitative metrics of teaching quality should be supplemented by a qualitative review of some sort.

5.4 Summary

Systems of measurement (where discussed) tend to draw on a basket of measures, due to the wide range of potential metrics. Quantitative metrics can be grouped into measures of inputs and outputs, process measures and rankings. Many authors are dubious about the value of key input (prices, fees, staffing numbers) and output measures (degrees awarded, jobs following graduation, raw student satisfaction scores).

There is more support for process measures looking at student engagement (though not systematically studied), but without clearly tried and tested tools for use in England. There is some evidence that students favour student engagement surveys over satisfaction surveys, and may take them more seriously.

Ranking metrics are generally regarded with caution, as the most respected rankings are based more on research excellence rather than any consideration of teaching excellence. There is little evidence in the literature on qualitative metrics, although a number of authors comment on the desirability of quantitative metrics of teaching quality being supported by some form of peer review, accreditation, visit or audit.

6. Student Focus Groups

6.1 Introduction

Fourteen focus groups were conducted in ten universities (5 pre-1992, 5 post 1992), involving 81 students. The subject areas included: Adult Nursing, Biochemistry, Biomedical Sciences, Business and IT, Business Information Systems, Chemistry, Children's Nursing, Criminology, Digital Media and Computing, Film, TV, and Digital Production and Sociology, Digital Media and Computing, Economics, History, Management with Information Systems, Marketing, Mathematics, Media Studies, Medical Genetics, Medicine, Politics and International Relations, Pharmacy, Physics, Physiotherapy, Psychology, Social Work.

The focus groups explored students' views in relation to the factors they considered when applying to university, the proposed teaching excellence framework and their current experiences of teaching at university. The universities included pre and post-1992 institutions, research-intensive and teaching-focussed universities and were in a situated in different parts of the country.

6.2 Applying to university

Factors students considered when deciding where to apply to

Whilst several students in two of the universities said teaching quality had been an important consideration when they applied to university, course structure and coverage were generally felt to be more relevant. Those students who had considered teaching quality said that this had involved using departments' research profiles and the profiles of academic staff as an indicator of teaching quality, and/or gauging teaching quality when attending open days or through use of the National Student Survey and the Unistats website or comparing UK higher education course data.

It was recognised across the focus groups that some information indicative of teaching quality was provided by the university, for example on membership of professional teachers' organisations. Several students also referred to meeting lecturers at open days as a means of learning about teaching quality, saying that they had been impressed by their enthusiasm and commitment. In the case of health and social work courses, the practice background of most lecturers was regarded in a positive light. With sandwich courses, which involve undergraduate students spending a period of placement with an employer as part of their degree course, the quality of placements was an important factor.

The fact that most students had not considered teaching quality when applying to university is not to say that this would have been the case had more information on teaching quality been readily available and/or students' had not assumed that they could rely on universities providing high quality teaching having, in some cases, used the reputations of universities, departments and staff as proxies for teaching quality.

The main factors students at pre-1992 institutions had considered when applying to university were: university reputation, reputation of department; and research profile of the department. Other factors mentioned were 'feel' of the university ("can you see yourself

living here?"), course content; and location. The factor most commonly mentioned by students in workshops at post-1992 universities was the desire to study at an institution that was local to them (described as "within commuting distance") in order to accommodate family commitments and/or financial constraints. This was also mentioned by students in workshops at two of the pre-1992 universities. One participant noted: "I was considering [another HEI] but there was no way I would be able to afford to live there".

Other factors mentioned by students at post-1992 universities included: the reputation of the course; the relevance of the degree course to securing employment post-qualification (including the availability of a placement year in industry within the course); positive "word of mouth" from friends and family; placement opportunities and links with industry; and a desire to study at a smaller institution, where the student would be "not a number but a name". A number of students had been impressed at the open days and interviews, feeling comfortable with the university, the mentors and the lecturers. The degree of support provided to students was also valued. Some students also mentioned employability outcomes as an important factor.

In one focus group at a post-1992 university, the majority of participants stated that they either applied to this HEI only (or to only a very small number of HEIs, as their key focus was on accessing an 'extended degree', (degrees that allow students without traditional A- level qualifications to access a staged degree programme). One participant stated that: "I couldn't have gone to university without it".

"When I was looking at the course I wanted to go into, the teaching quality was quite important. I went to an open day, and how the lecturers responded to questions – asking about teaching quality, about what I would receive as part of my course was quite important to me."

"Teaching quality should really be everyone's main factor when they are applying, but in the real world it never is – you picture the university you want to be at and look for that"

I don't think it [Teaching Quality] crossed my mind: you just assumed the teachers would be good. I think, if you're teaching at university level, you should be good."

"I focus on employability – the money we're paying, you want a job at the end of it. The actual teaching quality - most students would naturally assume it would be up to standard. I think what would swing it would be the opportunities you would receive, rather than the teaching quality because I don't think there'd be much difference between each uni. If you were wanting to move away [this student had wanted to remain in their local area] you'd probably look at the rankings, to look at the reputation of the course you wanted."

Sources of information accessed by students

The students had looked at universities prospectuses online, along with other websites – such as Which University? – which provided information on universities, league tables which rank universities, employment outcome information on the UCAS website (most, but not all, had looked at the UCAS rankings), and social media.

As noted above, many students had attended open days to gain more information on the university and, together with interviews and one-to-one sessions with academic staff, these were highlighted as being particularly valued by many of the students. They were viewed as especially important by those who were looking for information about the support universities provided to more "atypical" students, such as mature students and people with disabilities. Nonetheless, it was generally felt that that information provided by HEIs would be likely to "emphasise the positive".

There was agreement across the focus groups that judging the 'feel' was more important and that this was carried out "through unofficial means which give you a better view", such as student forums online and Facebook (which offer student's comments about their experiences at universities). For many students, these were the most useful and trusted sources of information. Informal conversations with friends and relatives were also mentioned.

In one focus group there was agreement that most universities prioritised undergraduate study over postgraduate study and that finding "any information about the reality of a postgrad course is rare".

"I spoke to student ambassadors when I came to the open day, and people who were previous students on the same course. I think it's a good recommendation if students have a good word to say about it. I know they're primed, they're not going to pick people who are going to slate them, but it is good to get information like that. They're still willing to say 'you won't like this bit of the course but you've got to do it."

"I had a family member who was doing a very similar course to mine in recent years, so that gave me better knowledge about the course. Also, I came to the open days, and looked at the reviews of past students. So I combined all these bits of information. (I also looked at the UCAS rankings)."

"I did look online, but it was mainly going to the open days and talking to other students."

The availability of reliable information on employment outcomes

At one focus one group participants recalled only a table in the prospectus that set out "types of work" or "career paths" for graduates from this HEI which they had found helpful. More generally students noted that employment outcome information was provided in the prospectus, at open days and (in some cases) it was mentioned at interview. Those studying medical and health related subjects noted that in their case it tends to be assumed that jobs will be available.

In general, students were sceptical about information provided by HEIs on employment outcomes, and thought that they "have to be taken with a very large pinch of salt". In one focus group, for example, this information was said to be easy to find on their university's website but it was felt that "it could not be trusted." In another focus group, students' points of scepticism included: whether the secured employment cited by their university was actually directly relevant to the degree studied; and whether the employment secured was actually what the graduate had wanted to pursue when they applied for their degree. Participants in two focus groups suggested that employment data and case studies

presented at open days were "done to impress the parents rather than the [prospective] student. Two participants in one group felt more useful data was gained from looking at the graduate recruitment information provided by large employers and "taking notice of which university[ies] they seem to prefer".

The availability of reliable information on the *A-level / qualifications* needed to best secure a place at the institution and course students were interested in

There was general agreement amongst the students attending focus groups that information on entry requirements was easy to find and clear. In one focus group, however, four participants held overseas or non A-Level qualifications, and applying had been "very confusing", with all participants reporting that different publications from the HEI had given different advice. Moreover, the case of two universities, participants said that it was made clear at open days that there would be flexibility on entry requirements. Two students also noted that one HEI had made it clear that there was a certain "leniency" in grade offers, and that a prospective student that fell slightly short of their offer would be able to enter into "dialogue" with the HEI.

Students who had started on foundation programmes, said that information on the process of transfer within from the foundation to main elements of their extended degrees was more relevant to them, and it was important that this had been easy to find and "very clear", especially in terms of identifying where criteria and procedures for the extended degrees differed between subjects. In one case, it was noted that at this HEI entry requirements are "the first thing that pop up" on the webpages for specific programmes.

The extent to which the quality of teaching has met students' expectations

Mixed views were expressed within and across the focus groups. Some students felt that their expectations had been exceeded (e.g., highlighting academic engagement, the 'passion' and commitment of the academic staff, and becoming involved in wider opportunities - such as contributing to research programmes); others felt that there expectations had been met; while still others highlighted "poor teaching styles" of some academic staff and programmes that were "not what was advertised at all". Several students commented that the teaching style of some lectures made it difficult to follow without some previous exposure to the subject. Overall, most were generally satisfied with the quality of teaching.

"It's a very high standard for my course. I think students now have to look more closely at what they're getting because now's there's a figure on what they have to pay to get the service. . . . But I have no complaints about how, and what, we are being taught. . . I don't think there's many on my course that could complain. I'm the course rep so I have more of an overview, and the satisfaction is high. The students really enjoy it, and the teaching quality's good."

"The cost of the course always influences your opinions . . . you're shelling out a lot of money and you're thinking 'is it really worth it?" The teaching's what I expected but at the level we're at now, at university, you really wouldn't expect it to be any less."

"It's exceeded my expectations because, at A levels, the teachers drilled it in that 'it's not like school – they don't spoon feed you' . . . I expected to come here and be thrown in the deep end, with no support. Actually, if I need support, if I need help on an assignment, they'll talk me through it. There are lectures on the assignments, helping you structure it. There's so much more support within the teaching than I expected."

"I leave some lectures thinking 'why on earth are you in teaching - you clearly hate doing it? […] Lecturers are always telling us that their door is open for support, but you go through some doors and there's no support on the other side".

"The teaching styles of some of our lecturers are poor not what was advertised at all". "You get

poor lecturers, but you have to get through it"

"Like everywhere I suppose, different teachers have different approaches and for the most part they suit my learning approach but one or two didn't. But the resources were there so, even if you don't exactly get on with the way a certain lecturer lectures, you can still do the work."

A relatively small number of students felt that degree subjects (and the mode of teaching those subjects) were so different to their experiences of pre-HE teaching that a comparison could not be made and/or that teaching quality had not been a factor they had considered, and as such, they had no expectations.

6.3 How TEF may affect the students' behaviour and decisions

How useful do students think the TEF would have been when they were applying to university?

The focus group facilitators explained that there is a proposal to introduce a new set of incentives whereby experts will provide a rating of the quality of teaching at different universities based on student satisfaction, student retention rates and graduate job prospects. The facilitators told the students that this is to ensure that HEIs are measured according to the outcomes and progress their students make. When students were asked how useful they thought this information on teaching quality would have been when they were applying to university, most students were agreed that they would have considered TEF data when applying to university (apart from those who were set on applying to a particular institution). However, many doubted whether TEF would have influenced their decision. The students were, in general, not very enthusiastic about a TEF at the institution level, although there were those who thought it might be of some value. Some students in post-1992 universities were sceptical of the value of such ratings because they felt "they always favour elite institutions". Some students felt that employment outcomes within the TEF would not be appropriate in their view, feeling that "graduate employability is not a measure of teaching quality – it just doesn't work like that [on my programme]".

"If it was in the public sphere, and if it was advertised, but there's no way I would have sought it out".

"It might (have been) useful if you have a choice between two or three universities".

"The more information you have – especially if it's non-biased – the better. Especially if it's set out clearly on each university"

"It would be good if there was a standardised rating that looks at all universities on an equal level. There are a number of ratings out there but they all rate them on different things, or use different values. (But) would it just become another table with Cambridge and Oxford at the top? Whenever I looked at a league table I thought 'this is fantastic – now scroll".

"I don't think student satisfaction surveys are reliable. I don't think the NSS has even been reliable for what it's supposed to do because the questions are very vague . . . The NSS favours the larger institutions, so it can be taken with a pinch of salt. . . . Student perception is not factual, it's based on opinions."

"My concern would be that, if so much emphasis were placed on figures, is that universities would end up teaching to get those figures rather than teaching what businesses are looking for in graduates. You would get a shift back to academic success rather than employment. . . A lot of the data is good initially, just to get a feel for what you're looking at but it's your decision, and it can be right or wrong for other students but it has to be right for you."

"This change is not going to make the lecturers better teachers"

"[The TEF] will lead to an 'infinite loop' where the top universities will always remain on the top. It makes me really angry".

"How useful would subject level information have been when students were applying to university?

The focus group facilitators explained that In the future, it is possible that the rating of teaching quality will be available at a discipline/subject level as well as institutional level. When students were asked how useful they thought this would have been when they were applying to university, there was broad agreement that a subject level TEF would be more useful than institutional level assessments. It was recognised that there might be a difference in approach between different departments in a university so it was necessary to focus on the specific subject you are interested in. TEF ratings and data would only be meaningful if "subject specific". However, a number also felt that other factors such as location, the "feel" of the place, how interviews had been conducted would have remained the major factors determining their choice. Some felt that their experience of interviews at some other universities had definitely deterred them from wanting to go there, and that a positive interview experience at their present HEI had been a key factor in their choice. Two students in the same focus group recalled that they had been actively discouraged by A Level teachers from studying their chosen degree subject and suggested that a subject level rating comparable with other subjects would have helped challenge this discouragement. Some participants thought that, rather than a subject level TEF, a better option would be for departments to publish their in-house student feedback surveys.

"That would be really useful, because an overall ranking for the university isn't really indicative of whether your course is going to be good or not."

"If they could get it down to individual courses it would be really good. I had the luxury of only having to choose between three universities - I could visit them all . . ."

"It could be a useful tool but my only concern would be students would be saying I want to do this course because its got a 90% positive feedback rating, for example, when they should be saying I want to do this course because it's right for me. I suppose if you could compare courses within a university it could be of use I suppose. But it's more important to establish what jobs you could do after that course."

"I think the information could be handy but I think the employability focus is the key. Knowing what degree will get you the job you want, that's the best kind."

To what extent is the TEF likely to be useful to people applying to universities in the future?

Many students thought that TEF will be useful to prospective students. However, some students felt that the value of a TEF would be felt more by employers in an 'increasingly competitive [market] for graduate jobs.

An analysis of students securing appropriate employment after their degree would be most useful. Some students emphasised the importance of using both short-term and long-term as, students graduating with some degrees (e.g. an MPharm degree) would tend to secure their desired employment relatively quickly, whilst students graduating with others (e.g. a MChem degree) may take longer to secure their desired employment.

In three focus groups students stated that they opposed the introduction of a TEF. They felt that "everyone will use it… it's not fair and it will kill real education [and] put up the cost". A further comment was: "it will widen the gap between academic and non-academic young people… assessing some government's idea of teaching quality is not going to improve the chances for young people… (it's) counter productive".

Other comments include:

""To have that extra information about the teaching on different courses would be helpful because I know nothing about the subject my son is intending to study."

"The data could be a good indicator for all students but it depends on how you translate it to your own needs."

"(It would) definitely be useful. My brother and sister applied this year – she had no idea what to look for, and she was always asking me what she should look at."

"People will start doing cost-benefit analyses and just not consider University… it's turning us into an American 2- or 3- tier system"

"It will widen the gap between academic and non-academic young people… assessing some government's idea of teaching quality is not going to improve the chances for young people… counter productive".

6.4 Current views on teaching quality within HE

Contact hours

When students were asked to what extent they felt they have a sufficient number of contact hours with the appropriate level of academic staff to support them in their studies, they expressed a mixture of viewpoints. Some felt that they had sufficient contact hours with appropriate academic staff. These students were in agreement that the contact hours varied between programmes, but generally felt they were appropriate for each programme. Other students were not content. Some said that they would welcome more contact on some of their courses, whilst others highlighted delays in academic staff responding to emails (or no response), and a lack of ability to book appointments with academic staff as key sources of frustration. Students in the humanities and social sciences generally expressed more concerns than did those in sciences and health subjects.

"We have a discussion board that our tutor checks, you can contact them that way as well if you need to."

"We have a Facebook group for each module, and they post stuff that we might find relevant. If we ask questions on it, they always answer quickly."

"Some lecturers are never available, but some are always there for you."

"I think some students have unrealistic ideas of when staff should be available. It's been brought up at staff-student meetings, when they've expected an email response within two hours."

Lecture and seminar class/tutor group sizes

While some students felt that their lecture and seminar/tutor group sizes were suitable for the needs of their course, others voiced concerns about this issue, with one participant feeling that "we were packed in like sardines" in some sessions in their first year (these were common sessions attended by students from a number of different degrees). However, several focus groups students observed that the group sizes had improved as they had progressed through their courses.

It two focus groups student attendance was identified as a "major problem". It was suggested that low attendance rates for some modules might be because a lot of the information transmitted in lectures is placed on electronic blackboards on the web. If students have assignments to complete they may avoid lectures and find what they need online.

"There are 30-40 students on the course in my year, and that's pretty good going. I know of courses where there's 200 odd students on one course and I could imagine that would be frustrating for some students. What's really good is peer monitoring – students can tell

you whether it's great or not. Academics can kind of disguise it, whereas students are very clear about what they like and don't like."

"we were packed in like sardines"

To what extent are students academically stretched and challenged?

The students generally felt stretched and challenged. In some cases students said that this varied by modules within their course, some feeling that they may, on occasion, be too stretched on some modules, and in others, not stretched enough. Moreover, in one focus group, there was general agreement that "some students are lazy – I know people that are choosing their options based on the reputation of how hard the modules are".

Students mentioned a difficulty with having the same deadline date for more than one assignment. It was also said that the timing of lectures was not always helpful, some being delivered close to assignment deadlines when it would have been helpful to have them earlier in the year. Some of the lead times, such as a period of four weeks to develop a proposal, were seen as too short.

"Regarding how challenging each particular module is, I think it's more important that it's relevant to what you want you want to do on graduation. If a module is a bit easier, as long as it's giving you the skills, I don't think it matters too much. If you were to do a really challenging module that was then redundant on graduation, there's not too much point to that."

" It gets to the point sometimes where I do feel I'm being stretched but the rest of them aren't on the same level so I feel I'm at school again, spending ages and ages on one single thing, so I get distracted. I have a few modules that I find more challenging than others but there's one particular module that I enjoy and want to get on but, every seminar, it's so slow."

To what extent do students feel engaged in your learning and motivated to pursue self-study?

Most students felt that they were engaged in their learning but some noted that levels of engagement were influenced by the extent to which lecturers were engaged in the learning process. With a handful of exceptions, students were in agreement that self-study was a key aspect of their programmes.

"It depends on how the lecturers encourage you. If they're not interested in the module then the students aren't encouraged . . . It's two way – if you see the lecturer is interested then you're interested. And if there's a connection with the work you'll be doing . . ."

Do you feel that your course is helping you develop the knowledge and skills you need for the future?

The students largely felt that their courses were helping them to develop the knowledge and skills they would need in the future. This was especially so in the case of students whose courses involved placements linked to employment after graduation.

"I feel I do on my course, the course we do is really industry-oriented – they adjust modules each year to what industry is looking for. I appreciate that other courses may not have that outlook, they may be more academic. But my course is excellent in looking at what we can do once we finish the course."

"We do a lot of group work, so that's useful, and team working skills."

Do you feel that you are receiving adequate careers guidance?

Students' perceptions of the visibility of the careers services and provision of career guidance varied. The HEI Careers Services were said in some cases to be reactive rather than proactive, and that "you have to have a set of career ideas already before they will help you". It was felt that a great deal of the guidance offered was at a basic level. Alongside a basic level of careers guidance, there was a large amount of guidance on "postgraduate courses rather than proper careers guidance". This was attributed to academic staff "giving advice on what they know about: post-graduate courses, postdoc research etc". One student said that careers advice could be "patronising" for mature students. Careers fairs were universally praised as being "self-directed: you can talk as much as you want about what you want".

To what extent do you receive feedback in a timely and helpful manner? What (if any) changes could be made to the way you receive feedback to make it more meaningful/useable?

There were examples of students feeling that feedback on their courses was poor. Some face-to - face feedback led students "coming out knowing less than I did before" and "just having to guess [what was required]". Others felt that the content of feedback was helpful, but that the timeliness was often very poor. Others felt that feedback was both timely and helpful. It was a mixed picture. The most commonly held view was that feedback was "variable", and that "the more you ask for, the more you get". One participant felt that this was appropriate as at degree level students should take an active role and "shouldn't be spoon [fed]".

"About 60% of the time": "It's generally good, but not always": "It's not always timely. It's supposed to be within three weeks "

"You have to appreciate they have a lot of work to mark and you can't expect it to be instant. But once you get it back it's good to read through, and I've found it to be effective. If there's an area where you've been marked down the reasons for that are clear."

Have you ever considered dropping out? What support have you received to help you stay on?

Several students in different groups had had considered dropping out and several said they had benefited from support from their personal tutor and institutions, although others did not seek support, "preferring to "deal with it on my own". One participant, for example, had been unhappy with their original programme and had been supported to swap programmes at the start of their second year. Several said that they had considered doing so, and that stress levels could be high. There was mention of "second year blues". But most felt that their HEIs provided support in such circumstances, helping students deal with such problems if they asked for support. Other reasons for considering dropping out included ill-health and bereavement; in these cases the students said the support that they had received had been excellent. It was also suggested by students that academic staff were very attuned to spotting and addressing declines in student attendance or progress.

"I've been here one semester. I'm really glad I'm going home at the weekend for two weeks, These past five weeks have been stressful, having to do school placements on Thursdays and Fridays, then a part-time job, then here at the university or three days a week. It has been ridiculously stressful, but . . .There's a girl on my course who is leaving as she's decided that the course isn't for her. Speaking to her about why she feels that way, I don't feel that way."

Overall, how satisfied are you with the quality of your course? Do you feel like you are getting value for money for your course? Why/ why not?

Some students were generally happy with both the overall quality and value for money of their programmes. Overall, the positives were thought to outweigh the negatives and that the fees bought "not just the course, but the whole experience". One student summed up the views of many: "I could learn this course [content] online for free… it's the networking, experiences, interactions… can't put a price on that".

Others were barely satisfied, citing lack of resources and poor lecturing styles. These students, who were based in three of the universities (including both pre and post 1992) did not feel that their programmes offered value for money', e.g. one observed that "I don't see how paying £9000 to watch someone read out a power point can be value for money", and "it's just warped – on some courses it's equivalent to something like £75 every time you see a lecturer or get an email from them… just wrong".

Participants discussed and agreed that the HEI sometimes "struggled to balance" academic aspirations with "the degree business". This struggle could on occasion be "heavy handed", with some lecturers "continually bigging-up their subject; it can sound like they are justifying the course to us and themselves – but we've already bought it!"

"It depends on my mood. If I'm just reading off a PowerPoint then what's the point – paying all that money just to see something written on a screen. But on other days, when they're a bit more passionate about it, when there's lots of stuff going on that you can get involved in, then it's good."

6.5 Summary

Teaching quality did not feature as a key issue that the students explicitly considered when applying to university. However this is not to say that would have been the case had more information been available- indeed, some students appeared to have used the reputations of universities, departments and academic staff as proxies for teaching quality. Issues students considered included university reputation, reputation of department; research profile of the department; and whether an institution was within commuting distance. Other factors mentioned were course content, the general 'feel' of the university ("can you see yourself living here?"), location, personal experiences and observations at open days and interviews, universities prospectuses online, other websites – such as Which University? – which provided information on universities, league tables which rank universities, employment outcome information on the UCAS website, and social media. For many students, fellow students' comments about their experiences at universities on social media websites such as Face Book and The Student Room were the most useful and trusted sources of information..

Most students were agreed that they would have considered TEF data when applying to university (apart from those who were set on applying to a particular institution), but many doubted whether TEF would have influenced their decision. The students were, in general, not very enthusiastic about a TEF at the institution level, although there were those who thought it might be of some value. Some students in post-1992 universities were sceptical of the value of such ratings because they felt "they always favour elite institutions". It was suggested that the TEF should be conducted at subject level due to differences in approaches to teaching in different disciplines/subjects. Some students were sceptical about the inclusion of employment outcomes within the TEF because they felt "graduate employability is not a measure of teaching quality; however, some thought data on the extent to which graduates secure appropriate employment after their degree would be useful (whether as part of the TEF, or alongside the TEF). Some students felt that the value of a TEF would be felt more by employers in an 'increasingly competitive [market] for graduate jobs.

Whilst most students were generally happy with both the overall quality and value for money of their programmes, some were not satisfied, citing lack of resources, poor lecturing styles, unhelpful or untimely feedback, and too few contact hours with appropriate staff. The latter are distributed across all of the institutions involved in the study.

Lecture and seminar class/tutor group sizes were a common concern, however. Interestingly, a large majority of students said that they were academically stretched and challenged and were engaged and motivated to pursue self-study. Students expressed mixed views on whether their courses were helping them develop the knowledge and skills they need for the future.

Some felt that the careers services and provision of career guidance could be more visible, especially during the first two years of degree programmes.

7. Issues to consider

The lack of agreement in the literature as to what constitutes teaching excellence in higher education is striking. Comparative research looking at perceptions of teaching excellence across different disciplines, universities and countries is needed. However, this is unlikely to be an easy task, as just as definitions of quality change over time and from one context to another, so too do definitions of excellence.

Future research on teaching quality in HE should take on board the complex relationships between proxies for teaching quality and actual teaching quality. Particular attention should be paid to how definitions of effective teaching evolve over time due to changing student needs, new modes of delivery, changing employer needs and technological innovation.

There is relatively little research into academics' understanding and perceptions of teaching excellence and further research in this area is sorely needed. It would be particularly helpful to look at how academics and students from the same institutions and courses perceive teaching quality and excellence, and also how these perceptions may change over time. There is also the issue of shifting thresholds, and how these can be accommodated in systems designed to measure excellence.

There is a relative paucity of literature on the potential for using types of learning engagement as a metric in a set aimed at identifying learning excellence. This may warrant further detailed investigation as a separate research topic when developing a potential battery of metrics for teaching excellence.

There is no evidence on how teaching quality metrics can be broken down to different groups of learners, particularly groups differentiated by learner characteristics rather than subject/programme studied. The metrics that might pick that up – student engagement – are measured on voluntary completion of survey instruments, which may not be applied to all final year undergraduates in an institution, and will certainly not be completed by all final year undergraduates. Survey instruments are in any case meant to be anonymous. Disadvantage and under-representation will also be specific to programme and to institution, depending on catchment, fees and other characteristics.

Neither the benefits accruing to HEIs from an increase in teaching quality nor the views of employers regarding teaching quality in HE have been systematically explored in the UK. Research is urgently required in both of these areas.

A rich body of research is emerging in relation to the Undergraduate Teaching Evaluation (UTE) scheme in China, which provides insights that may be relevant in England, especially in relation to the use of student engagement and course evaluation surveys in national HE teaching quality assessment systems.

References

Aleamoni, L. M. (1999). Student rating myths versus research facts from 1924–1998. *Journal of Personnel Evaluation in Education*, 13: 153–166.

Altbach, P. (2015b). *The Dilemmas of Ranking, International Higher Education*. http://ejournals.bc.edu/ojs/index.php/ihe/article/viewFile/7878/7029

Alderman, G. (1996). Quality assessment in England: a worm in the bud? In T. W. Banja and J. L. Lambert (eds.) *Proceedings of the Eighth International Conference on Assessing Quality in Higher Education, July 14-16, 1996*. Indiana: Indiana University.

Ashwin, P., Abbas, A., McLean, M. (2014). How do students' accounts of sociology change over the course of their undergraduate degrees? Higher Education, 67, 2, 219 – 234.

Australian Learning and Teaching Council. (2008). Australian awards for university teaching: Guidelines for 2008. webdav/site/carricksite/users/siteadmin/public/2006%20Awards%20Guidelines.pdf

Australian Learning and Teaching Council. (2009). History of the ALTC. http://www.altc.edu.au/carrick/go/home/about/pid/570

Ball, S. (2003). The teacher's soul and the terrors of performativity." *Journal of Education Policy*, 18 (2): 215–228.

Barre, E, (2015), Do student evaluations of teaching really get an "F"?, Blog, Rice Centre for Teaching.Excellence,http://cte.rice.edu/blogarchive/2015/07/09/studentevaluations

Barnett, R. (2011). *Being a University*. London and New York: Routledge.

Barnett, R., Parry , G., Cox , R., Loder C. & Williams , G. (1994). *Assessment of the Quality of Higher Education*. Report for HEFCE and HEFCW (London, University of London, Centre for Higher Education Studies, Institute of Education).

Benson, R., & Samarawickrema, G. (2009). Addressing the context of e-learning using transactional distance theory to inform design. *Distance Education*, 30(1): 5–22

Berk, R. (2005). Survey of 12 strategies to measure teaching effectiveness. *International Journal of Teaching and Learning in Higher Education*,17(1): 48 – 62

Bornmann, L., Mittag, S. and Daniel, D. (2006). Quality assurance in higher education – meta-evaluation of multi-stage evaluation procedures in Germany. *Higher Education*, 52: 687–709. DOI 10.1007/s10734-004-8306-0

Bologna Declaration (1999). The Bologna Declaration of 19 June 1999. Bologna: Joint Declaration of The European Ministers of Education.

Bouchard, J. (2013). Les classements d'établissements d'enseignement supérieur et de recherche : des miroirs déformants aux instruments de régulation, Questions de communication, 2013/1 (n° 23), http://www.cairn.info/revue-questions-de-communication- 2013-1-page-175.htm

Bovill, C. (2014). An investigation of co-created curricula within higher education in the UK, Ireland and the USA. *Innovations in Education and Teaching International* (In Press).

Bradley, S., E. Kirby, and M. Madriaga (2015). What students value as inspirational and transformative teaching. *Innovations in Education and Teaching International*, 52(3): 231– 242.

Brennan, J., and T. Shah (2000). *Managing Quality in Higher Education: An International Perspective on Institutional Assessment and Change*. Society for Research into Higher Education, Buckingham: Open University Press.

Brown, G., & Atkins, M. (1993). *Effective teaching in higher education*. London: Routledge. Brown,

R. (2013). Mutuality meets the market: Analysing changes in the control of quality assurance in United Kingdom higher education 1992-2012. *Higher Education Quarterly*, 67(4): 420-437. DOI: 10.1111/hequ.12028.

Browne Review. (2010). *Securing a Sustainable Future for Higher Education*. Independent review. Retrieved April 16, 2013, from www.independent.gov.uk/browne-report

Brusoni, M. (et al.) (eds.) (2014). *The Concept of Excellence in Higher Education.*Brussels: European Association for Quality Assurance in Higher Education

Buckley, A (2014) UK Engagement Survey 2014 The second pilot year. York: Higher Education Academy

Cardoso, S., Rosa, M.J. and Stensaker, B. (2015). Why is quality in higher education not achieved? The view of academics. *Assessment and Evaluation in Higher Education*. DOI: 10.1080/02602938.2015.1052775

Carnell, E. (2007). Conceptions of effective teaching in higher education: Extending the boundaries. Teaching in Higher Education, 12(1) 25–40.

Carr, S., E. Hamilton, and P. Mead. 2005. Is it possible? Investigating the influence of external quality audit on university performance. Quality in Higher Education 11, no. 3: 195– 211

Chaffee, E. (2014). Learning metrics: How can we know that students know what they are supposed to know? *Association of Governing Boards of Universities and Colleges, Trustees Magazine*, January/February 2014, http://agb.org/trusteeship/2014/1/learning- metrics-how-can-we-know-students-know-what-they-are-supposed-know

Chalmers, D. (2008). Teaching and learning quality indicators in Australian Universities. OECD/IMHE. Conference Outcomes of Higher Education: Quality, relevance and impact, Paris. http://www.oecd.org/site/eduimhe08/41216416.pdf

Chen, Y. (2008). Evaluation of undergraduate teaching at institutions of higher education in China: Problems and reform. *Fudan Education Forum*, 5(2): 5–7, 29 (in Chinese).

Cheng, M. (2010). Audit cultures and quality assurance mechanisms in England: a study of their perceived impact on the work of academics. *Teaching in Higher Education*, 15(3): 259-271. DOI: 10.1080/13562511003740817

Cheng, M. (2011). The perceived impact of quality audit on the work of academics. *Higher Education Research & Development*, 30(2): 179-191. DOI:10.1080/07294360.2010.509764

Cook, R., Butcher, I. and Raeside, R. (2006). Recounting the scores: An analysis of the QAA Subject Review grades 1995-2001. *Quality in Higher Education*, 12(2): 135-144.

Cuseo, J. (undated). http://www.classsizematters.org/wp-content/uploads/2012/11/Week- 13-Cuseo-1.pdf

Darian, L. (2016). Designing a Teaching Excellence Framework: Lessons from Other Sectors, Higher Education Policy Institute, Working Paper 13.

Davies, R., Hope, M. & Robertson, A. (2012) Student-Led Teaching Awards: lessons from a leading higher education initiative in Scotland. Higher Education Academy Scotland: Edinburgh.

http://www.heacademy.ac.uk/assets/documents/nations/scotland/HEA_SLTA_WEB_151112_1010.pdf

Dearing, R. (1997). The National Committee of Inquiry into Higher Education (NCIHE). London: Department for Education and Skills: p.7.

De Courcey, E. (2015). Defining and Measuring Teaching Excellence in Higher Education in the 21st Century. *College Quarterly* 18 (1). http://collegequarterly.ca/2015-vol18-num01-winter/decourcy.html

Department for Business, Innovation & Skills (BIS) (2011) *Students at the Heart of the System* http://bis.gov.uk/assets/biscore/higher-education/docs/h/11-944-higher-education- students-at-heart-of-system.pdf.

Devlin, M. (2007a, August). Improving teaching in tertiary education: Institutional and individual influences. Keynote address at Excellence in Education and Training Convention, Singapore Polytechnic, Singapore.

Devlin, M. (2007b). An examination of a solution-focused approach to university teaching development. Unpublished doctoral dissertation, The University of Melbourne, Australia. Downloaded by [Durham University Library] at 06:21 12 March 2016 Higher Education Research & Development 123

Devlin, M. (2007c, May). The scholarship of teaching in Australian higher education: A national imperative. Keynote address at the Vice-Chancellors Colloquium, The University of the Sunshine Coast, Queensland, Australia.

Dill, D.D. (2000). Capacity building as an instrument of institutional reform: Improving the quality of higher education through academic audits in the UK, New Zealand, Sweden, and Hong Kong. *Comparative Policy Analysis*, 2(2): 211–34.

Drennan, L.T. (nd) Teaching quality assessment scores and their influences – The Scottish experience.
http://www.cem.org/attachments/ebe/Lynn%20Drennan%20-%20Paper.pdf

Drennan, L. and Beck, Mathias (2001). Teaching quality assessment scores: measuring quality or confirming hierarchy? The Sixth QHE Seminar: The End of Quality? Birmingham, 25–26 May, 2001. http://www.qualityresearchinternational.com/papers/drennanpmv6.pdf

Elton, L. (1998). Dimensions of excellence in university teaching. *International Journal for Academic Development*, 3(1): 3–11.

Endrizzi, L. (2014). *La qualité de l'enseignement : un engagement des établissements, avec les étudiants ?* Dossier de veille de l'IFÉ, n° 93, juin. Lyon : ENS de Lyon. En ligne : http://ife.ens-lyon.fr/vst/DA/detailsDossier.php?parent=accueil&dossier=93&lang=fr

Fanghanel, J., McGowan, S., Parker, P., Mcconnell, C., Potter, J., Locke, W., Healey, M. (2015) Literature Review. Defining and supporting the Scholarship of Teaching and Learning (SoTL): a sector-wide study. York: Higher Education Academy

Feigenbaum, A. and Iqani, M (2015). Quality after cuts? Higher education practitioners' accounts of systematic challenges to teaching quality in times of 'austerity'. *Journal of Further and Higher Education*, 39(1): 46-66. DOI:10.1080/0309877X.2013.778961

Fitzpatrick, M., and S. Moore (2013). Exploring both positive and negative experiences associated with engaging in teaching awards in a higher education context." *Innovations in Education and Teaching International*, 52(6): 621–631. doi.org/10.1080/14703297.2013.866050.

Forstenzer, J., (2016). The Teaching Excellence Framework: What's the Purpose? The University of Sheffield, The Sir Bernard Crick Centre, http://www.crickcentre.org/wp-content/uploads/2016/01/TEF-Whats-the-Purpose-booklet-Josh-Forstenzer.pdf

Foskett, N. Roberts, D. and Maringe, F. (2006). Report of a Higher Education Academy funded research project 2005-2006. [online] Available at: http//www.highereducationacademy. ac.uk.4407.htm

Gärdebo, J. & Wiggberg, M. (Eds.) (2012). Students, the University's Unspent Resource: Revolutionising Higher Education Through Active Student Participation. Uppsala Universitet. http://www.uadm.uu.se/upi/arkiv/rapporter/Rapport12.pdf

Gaspard, J, 2013, Le discours promotionnel des universités européennes. Homogénéité dans la compétitivité ?, Mots. Les langages du politique 2013/2 (n° 102), http://www.cairn.info/revue-mots-2013-2-page-53.htm

Gerbic, P., and I. Kranenburg (2003). The impact of external approval processes on programme development. *Quality in Higher Education*, 9(2): 169–77

Gibbs, G. (2007). Have we lost the plot with teaching awards? *Academy Exchange*, 7: 40– 42.

Gibbs, G. (2012). *Implications of 'Dimensions of Quality' in a Market Environment*. HEA Research Series. York: Higher Education Academy.

Gibbs, G. (2010). Dimensions of Quality, York: Higher Education Academy

Gibbs, G, 2016, In HEPI, *Response to the Higher Education Green Paper 2016* (op cit). Gift, S.I.,

and C. Bell-Hutchinson (2007). Quality assurance and the imperatives for improved students experiences in higher education: The case of the university of the West Indies. *Quality in Higher Education* 13(2): 145–57.

Gillard, J. (2009). Speech at Universities Australia Conference, March 4th, 2009. Retrieved March 23, 2009, from http://www.deewr.gov.au/Ministers/Gillard/Media/Speeches/Pages/ Article_090304_155721.aspx

Goffman, E. 1959. *Presentation of Self in Everyday Life*. New York: Doubleday Anchor Books.

Gosling, D. and D'Andrea, V. (2001). Quality development: a new concept for higher education. Paper presented at The Sixth QHE Seminar, The End of Quality? Birmingham, 25–26 May, 2001.

http://www.qualityresearchinternational.com/papers/goslingv3.pdf

Grifoll, J. (2014). Excellence in higher education. In M. Brusoni (et al.) (eds.) *The Concept of Excellence in Higher Education.*Brussels: European Association for Quality Assurance in Higher Education, pp.20-21.

Grove, J., (2015). How might a teaching excellence framework be built? Times Higher Education, July 23

Grove, J, 2016, Universities 'most reliant' on teaching-only staff named, Times Higher Education, 15 March, https://www.timeshighereducation.com/news/universities-most- reliant-teaching-only-staff-named

Gunn, V. (2013). Mimetic desire and intersubjectivity in disciplinary cultures: constraints or enablers to learning in higher education? *Studies in Continuing Education*, DOI:10.1080/0158037X.2013.787981

Gunn, V., and Fisk, A. (2013). *Considering Teaching Excellence in Higher Education: 2007–2013: A Literature Review Since the CHERI Report 2007*. York: Higher Education Academy.

Gunn, V., (2016), currently unpublished research looking at direct, indirect and emerging regulatory frameworks affecting higher education.

Gynnild, V. (2007). Quality assurance reconsidered: A case study. *Quality in Higher Education*, 13 (3): 263–73.

Hammer, D., Piascik, P., Medina, M., Pittenger, A., Rose, R., Creekmore, F.,Soltis, R., Bouldin, A., Schwarz, L. & Scott, S.(2010) Recognition of teaching excellence. American Journal of Pharmaceutical Education, 74 (9), 1-11.

Hannan, A., & Silver, H. (2000). *Innovating in Higher Education: Teaching, Learning and Institutional Cultures*. Buckingham: Open University Press.

Harrison and Lockwood (2001).

Harvey, L. (2002) Evaluation for what? *Teaching in Higher Education*, 7(3): 245–263.

Harvey, L. (2005). A history and critique of quality evaluation in the UK. *Quality Assurance in Education*, 13(4): 263–276.

Harvey, L., and J. Newton. 2004. Transforming quality evaluation. *Quality in Higher Education*, 10 (2): 149–65.

Hénard, F and Roseveare, D (2012), Fostering Quality Teaching in Higher Education: Policies and Practices. OECD, An IMHE Guide for Higher Education Institutions http://search.oecd.org/edu/imhe/QT%20policies%20and%20practices.pdf Henkel, M. (2000). *Academic Identities and Policy Change in Higher Education*. London and Philadelphia: Jessica Kingsley.

Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1): 53–64.

Higher Education Policy Institute, (2016). *Response to the Higher Education Green Paper*. HEPI report 81.

Hill, Y., Lomas, L., & MacGregor, J. (2003). Students' perceptions of quality in higher education. Quality Assurance in Education, 11, 15–20.

Holroyd, J., and Saul, J. (2016). Will the Teaching Excellence Framework be Sexist?. The Guardian, 4 April 2016, http://www.theguardian.com/higher-education- network/2016/apr/04/will-the-teaching-excellence-framework-be-sexist

Horsburgh, M. (1997). External quality monitoring in New Zealand tertiary education. *Quality in Higher Education*, 3(1): 5–15.

James, R., Bexley, E., Devlin, M., & Marginson, S. (2007). *Australian University Student Finances 2006: Final report of a national survey of students in public universities*. Canberra: Australia Universities Australia.

Jiang, K. (2010). Undergraduate teaching evaluation in China: Progress and debate. *International Higher* Education, 58: 15–17.

Kandiko-Howson, C. B. & Mawer, M. (2013). *Student Expectations and Perceptions of Higher Education: Executive Summary.* London: King's Learning Institute.

Kandiko-Howson, C. B. (2016). Student Engagement and Teaching Excellence: Metrics for Enhancement. University of Greenwich Open Lectures in Teaching and Learning, 2015-16 Series, https://vimeo.com/154241189

King, R. (2011). *The Risks of Risk-Based Regulation: The Regulatory Challenges of the Higher Education White Paper for England*. Oxford: Higher Education Policy Institute.

Kogan, M., Bauer, M., Bleilie, I., & Henkel, M. (2000). *Transforming Higher Education: A Comparative Study*. London: Jessica Kinsley.

Kohoutek, J. 2009. "Quality Assurance in Higher Education: A Contentious Yet Intriguing Policy Issue." In Implementation of the Standards and Guidelines for Quality Assurance in Higher Education in the Central and East European Countries – Agenda Ahead, edited by J. Kohoutek, 21–50. Bucharest: UNESCO-CEPES.

Kohoutek, J. (2014). European standards for quality assurance and institutional practices of student assessment in the UK, the Netherlands and the Czech Republic." *Assessment & Evaluation in Higher Education*, 39(3): 310–325. doi:10.1080/02602938.2013.830694.

Kreber, C. (2002). Teaching excellence, teaching expertise, and the scholarship of teaching. Innovative *Higher Education* 27: 5–23.

Land, R and Gordon, G (2015) *Teaching Excellence Initiatives: Modalities and Operational Factors*. York: HEA.https://www.heacademy.ac.uk/resource/teaching-excellence- initiatives-modalities-and-operational-factors

Laughton, D. (2003). Why was the QAA approach to teaching quality assessment rejected by academics in UK HE? *Assessment and Evaluation in Higher Education*, 28(3): 309-321.

Lee, J. C. K., Y. X. Huang, and B. Zhong (2012). "Friend or foe: The impact of undergraduate teaching evaluation in China. *Higher Education Review*, 44(2): 5–25.

Little, B., Locke, W., Parker, J., and Richardson, J. (2007). *Excellence in Teaching and Learning: A Review of the Literature of the Higher Education Academy*. York: Higher Education Academy. DOI: 10.1080/13562517.2015.1136277

Little, B. and Locke, W. (2011), "Conceptions of excellence in teaching and learning and implications for future policy and practice", in Rostan, M. and Vaira, M. (Eds.) (2011), Questioning Excellence in Higher Education – Policies, Experiences and Challenges in National and Comparative Perspective, Sense Publishers, Rotterdam.

Liu, S. (2013). Quality assessment of undergraduate education in China: Impact on different universities. *Higher Education*, 66(4): 391–407.

Lowman, J., & Mathie, V. A. (1993). What should graduate teaching assistants know about teaching? *Teaching of Psychology*, 20(2), 84-88.

Lu, G. 2007. "Enhancing Students' Learning Experiences, Improving the Teaching Quality of Higher Education." Fudan Education Forum 5 (2): 8–11 (in Chinese).

Madriaga, M. and Morley, K. (2016). Awarding teaching excellence: 'what is it supposed to achieve?' Teacher perceptions of student-led awards, *Teaching in Higher Education*, 21:2, 166-174.

Macfarlane, B. (2007). Beyond performance in teaching excellence." In *International Perspectives on Teaching Excellence in Higher Education*, edited by Alan Skelton, 48–59. Abingdon: Routledge.

Mahoney, C. (2012, July). Opening address: Great expectations – Are you ready. Higher Education Academy Annual Conference, Manchester University. Retrieved October 4, 2012, from http://www.heacademy.ac.uk/news/detail/conference_2012_underway

Mandelson, P. (2009) Speech at the CBI Higher Education Summit, London. [online] Available at: http:// www.news.bbc.co.uk/1/hi/ education/8316658.stm.

Mandernach, B. J. (2015). Assessment of student engagement in higher education: A synthesis of literature and assessment tools. *International Journal of Learning, Teaching and Educational Research*, 12(2): 1-14.

Marsh, H.W. (1994). *Students' Evaluation of Educational Quality (SEEQ): A Multidimensional Rating Instrument of Students' Perceptions of Teaching Effectiveness*. Sydney: University of Western Sydney Self Research Centre.

Marsh, H.W., & Roche, L.A. (1994). *The Use of Students' Evaluations of University Teaching to Improve Teaching Effectiveness*. Canberra: Australian Government Publishing Service.

Martin, E., Benjamin, J., Prosser, M., & Trigwell, K. (1999). Scholarship of teaching: A study of the approaches of academic staff. In C. Rust (Ed.), Improving student learning: Improving student learning outcomes (pp. 326–331). Oxford: Oxford Centre for Staff Learning and Development, Oxford Brookes University.

Massy W.F. (2010). Metrics for Efficiency and Effectiveness in Higher Education: Completing the Completion Agenda, White Paper, in Front and Center, Critical Choices for Higher Education. The report is based on a Jointly Sponsored Meeting of the Miller Center, Association of Governing Boards of Universities and Colleges, and National Conference of State Legislatures, and National Governors Association, Center for Best Practices, Charlottesville, VA, December 5–6, 2010.

McGettigan, A. (2013). *The Great University Gamble*. London: Pluto Press. McCulloch, A.

(2009) The student as co-producer: learning from public administration
about the student university relationship. *Studies in Higher Education*, 34 (2):171-183.

Soilemetzidis, I., Bennett, P., Buckley, A., Hillman, N. and Stoakes, G (2014) The HEPI- HEA Student Academic Experience Survey 2014, York: Higher Education Academy.

Middlehurst, R. (1997). Quality enhancement for accountability and transformation: A framework for the future, *Tertiary Education and Management*, 3(1): 15–24.

Mohrman, K., Y. Wang, and Xu Li. (2011). Quality assurance in undergraduate education:Transformation of higher education policy in China." In The Impact and Transformation of Education Policy in China, edited by T. Huang and A. W. Wiseman, 345–375. Bingley: Emerald.

Moore, S. & Kuol, N. (2007). Retrospective insights on teaching: exploring teaching excellence through the eyes of the alumni. *Journal of Further and Higher Education*, 31:2,133-143

Morley, L. (2003). *Quality and Power in Higher Education*. Maidenhead, UK: SRHE & Open University Press.

Neave, G. (1988). On the cultivation of quality, efficiency and enterprise: An overview of recent trends in higher education in Western Europe, 1986–1988. *European Journal of Education*, 23 (1/2): 7–23.

Newton, J. (1999). An evaluation of the impact of external quality monitoring on a higher education college (1993-1998). *Assessment and Evaluation in Higher Education*, 24(2): 215-235

Newton, J. (2000). Feeding the beast or improving quality? Academics' perceptions of quality assurance and quality monitoring. *Quality in Higher Education*, 6(2):153–164.

Newton, J. 2002. Views from below: Academics coping with quality. *Quality in Higher Education*, 8 (1): 39–61.

Ory, J. C. (2001). Faculty thoughts and concerns about student ratings. *New Directions for Teaching and Learning,* 87: 3–15.

Palfreyman, D., (2015) Quality in Higher Education. Research Intelligence, Issue 128, Autumn 2015, BERA

Parpala, A., S. Lindblom-Ylänne, and H. Rytkönen (2011). Students' conceptions of good teaching in three different disciplines. *Assessment & Evaluation in Higher Education*, 36 (5): 549–563.

Pol, P, 2012, La passion de l'excellence dans l'enseignement supérieur en Allemagne, en Espagne et en France, Repères May 2012,
http://ressources.campusfrance.org/publi_institu/agence_cf/reperes/fr/reperes_14_fr.pdf

Ramsden, P. Margetson, D. Martin, E. and Clarke, E. (1995). *Recognising and Rewarding Good Teaching in Australian Higher Education*. Australian Government Publishing Service, Canberra.

Rosa, M. J., D. Tavares, and A. Amaral (2006). Institutional consequences of quality assessment. *Quality in Higher Education*, 12(2): 145–159.

Saarinen, T. (1995). Systematic higher education assessment and departmental impacts: Translating the effort to meet the needs. *Quality in Higher Education*, 1(3): 223–34.

Schleicher, A. (2016). Value Added: how do you measure whether universities are delivering for their students, Higher Education Policy Institute Annual Lecture 2015, Report 82, http://www.hepi.ac.uk/wp-content/uploads/2016/01/Andreas-Schleicher-lecture.pdf

Schwarz, S., and D. Westerheijden, eds. (2004). *Accreditation and Evaluation in the European Higher Education Area*. Dordrecht: Kluwer Academic.

Shah, M. (2012). Ten years of external quality audit in Australia: evaluating its effectiveness and success. *Assessment and Evaluation in Higher Education*, 37(6): 761- 772.

Silver, H. 1995. *The external examiner system: Possible futures*. A report of a project commissioned by the Higher Education Quality Council. London: Open University Quality Support Centre.

Skelton, A. (2002). Towards inclusive learning environments in higher education? Reflections on a professional development course for university lecturers. *Teaching in Higher Education*, 7(2): 193–214.

Skelton, A. (2007a) "Introduction." *In International Perspectives on Teaching Excellence in Higher Education*, edited by Alan Skelton, 1–12. Abingdon: Routledge.

Skelton, A. (2007b). The National Teaching Fellowship Scheme 2000-2006: Rest in Peace? In A. Skelton (Ed.), *International perspectives on teaching excellence in higher education* (pp. 213–225). Abingdon: Routledge.

Smeby, J.C., and B. Stensaker (1999). National quality assessment systems in the Nordic countries: Developing a balance between external and internal needs? *Higher Education Policy*, 12(1): 3–14

Srikanthan, G., & Dalrymple, J. (2003). Developing alternative perspectives for quality in higher education. *International Journal of Educational Management*, 17(3): 126–136.

Stensaker, B. Langfeldt, L., Harvey, L. Huisman, J and Westerheijden, D. (2011). An in- depth study on the impact of external quality assurance. *Assessment and Evaluation in Higher Education*, 36(4): 465-478.

Stensaker, B. (2014). European trends in quality assurance: New agendas beyond the search for convergence?" In *Quality Assurance in Higher Education*, edited by M. J. Rosa and A. Amaral, 135–148. Basingstoke: Palgrave MacMillan. Westerheijden, Hulpiau, and Waeytens 2007.

Strike, K. A. (1985). Is there a conflict between Equity and Excellence? *Educational Evaluation and Policy Analysis* 7(4): 409-416.

Swain, H, 2015, Universities say no to new ranking, The Guardian, http://www.theguardian.com/education/2015/jul/28/pisa-style-ratings-higher-education- hefce-learning-gain-students-university

Thomas S, Chie QT, Abraham M, and Raj, S.J. (2014). A qualitative review of literature on peer review of teaching in higher education: An application of the SWOT framework. *Review of educational Research*, 84 (1): 112-159. http://rer.sagepub.com/content/84/1/112.full.pdf+html

Thune, C. (1996). The alliance of accountability and improvement: The Danish experience. *Quality in Higher Education*, 2: 21–32.

Trigwell, K. (2001). Judging university teaching. *International Journal for Academic Development*, 6(1), 65–73.

Tsui, A. (2013) Re-imagining Undergraduate Education. Keynote address, Scottish Enhancement Themes Conferencehttp://www.enhancementthemes.ac.uk/enhancement- themes-conference/conference-resources/keynote-presentat

Vroeijenstijn, A.I. (1995). *Improvement and Accountability: Navigating between Scylla and Charybdis. Guide for External Quality Assessment in Higher Education (Higher Education Policy Series, 30)*. Melksham, Wiltshire: Cromwell Press.

Vroeijenstijn, A.I. (2000). External quality assessment (EQA) in the Netherlands. The 3rd generation 2000–2006'. In Hochschulrektorenkonferenz (ed.), Leitbild der Hochschule – Qualita¨t der Lehre (Beitra¨ ge zur Hochschulpolitik; 2/2000). Bonn: Hochschulrektorenkonferenz (HRK), Projekt Qualita¨ tssicherung, pp. 53–66.

Wächter, B. and Kelo, M. (2015). *University Quality Indicators: A Critical Assessment*. Study 04/2015, for the European Parliament's Committee on Culture and Education, http://www.europarl.europa.eu/studies

Wilsdon, J., et al. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. HEFCE, DOI: 10.13140/RG.2.1.4929.1363

Westerheijden, D.F. (2015). Compare universities your way: U-Multirank and its uses for institutional management. In: SGroup general assembly 2015, "Connecting the world by shaping a global mindset - the impact of internationalisation strategies at SGroup's universities", 06-11-2015, Las Palmas. Presentation slides - http://doc.utwente.nl/98304/1/UMR-SGroup%20GranCanaria%201511.pdf

Yelland, R. and Valle, R.C. (2013). Towards an international assessment of higher education learning outcomes: the OECD-led AHELO initiative, in *Rankings and Accountabiliy: Uses and Misuses*, (eds.) Marope, P, Wells P and Hazelkorn, E, UNESCO Publishing.

Yin, H., Lu, G. and Wang, W. (2014). Unmasking the teaching quality of higher education: students' course experiences and approaches to learning in China. *Assessment and Evaluation in Higher Education*. 39(8): 949-970.

Yin, H. and Wang, W. (2015). Assessing and improving the quality of undergraduate teaching in China: the Course Experience Questionnaire. *Assessment and Evaluation in Higher Education* 40(8) 1032-1049.

Yorke, M. (1996) Shouldn't quality be enhanced, rather than assessed? *Tertiary Education and Management*, 2(1): 86–94.

Yorke M., Orr S. and Blair B,. (2014). Hit by a perfect storm? Art & Design in the National Student Survey, *Studies in Higher Education*, 39(10): 1788-1810.

Zeitzeva, E., (2016) Investigation of relationship between student engagement and professional recognition of teaching standards (HEA Fellowships), York: Higher Education Academy, https://www.heacademy.ac.uk/sites/default/files/downloads/ukes_and_hea_fellowship_correlation_march_2016.pdf

Annex 1 - Metrics: Strengths and Weaknesses

This table summarises the strengths and weaknesses of potential metrics for measuring teaching quality in HE..

Nature of metric	Description	Strengths	Weaknesses
Input	Access to resources (Kandiko-Howson 2013, 2015)	Easy to verify if resources exist, in 2013 considered important, in 2015 seen more as pass/fail style factor	These are hygiene factors and need to be there, but don't make for excellence
Input	Price (Massey)	Easy to identify	In US, linked to status and administration of the University.
Input	Teaching portfolio (Berk)	This is not discussed in any detail in the literature other than being identified as a possible measure	Generally individual level
Input	Class size (Gibbs, Cuseo, Soilemetzidis et al, Kandiko-Howson 2013)	Is understandable, is clearly linked to the quality of the learning experience	Will be influenced by academic year, and nature of course, and will need some form of differentiation between lecture, tutorial and workshop sessions
Input	Contact hours (Soilemetzidis et al)	Planned hours are easily calculated, learning analytics may make actual hours more easily captured	Will be influenced by programme type, may also be affected by cost of teaching, and so might need to be linked to the qualifications or status of the person providing the contact hours
Input	Student/faculty ratio (Chaffee, Kandiko-Howson 2013, Grove)	Is easily calculated, is considered important by students, is included in a number of the international rankings	Clear metadata required to understand calculation method, could be regarded as a measure of the wealth of an institution.

Nature of metric	Description	Strengths	Weaknesses
Input	Student and faculty characteristics (Chaffee, Kandiko-Howson 2013, Grove)	Should be easy to develop and present. HESA already collects data, including staff qualifications. Teaching knowledge including proportion of staff recognised within the UK Professional Standards Framework should be available at department level	May be linked to University history and reputation. Some concerns in sector about teaching only posts and part time posts and how that would be reflected. Metadata will need to be clearly stated.
Input	HEA Fellowship among faculty members	Should be easy to calculate and present, needs to take account of fellowship level, is a proven link to improved student engagement so may work as a proxy for that	HEA fellowship may not be supported by some institutions
Input	L'Etudiant rating system (web, Bouchard)	Visual, based on experience	Voluntary, low numbers of ratings even on large courses
Input	Learning analytics (Richards)	Can offer measures of how a student engages with the learning opportunities available	Still early days, tool for supporting teaching, not for measuring excellence. Will gather information on activity, but not necessarily on affective engagement with the subject
Outputs			
Output	Teaching awards (Berk)	Public (often very public) accolade for individual teachers	Dependent on institution
Output	Alumni ratings (Berk, Chaffee)	Based on reflection following graduation	Likely to be challenging to implement, including tracking alumni
Output	Employer ratings (Berk, Chaffee)	Authors who mention this are American and it applies to the American system	No history in England, no track record of measuring, potential to be swayed by reputation.

Nature of metric	Description	Strengths	Weaknesses
Output	Retention and graduation rates (Chaffee, Gibbs, Ashwin)	Retention and completion rates should be useful and consistent measures	Need consistent ways to calculate measures
Output	Degree outcome (HEPI, Gibbs) Learning outcome measures (Berk, Kandiko-Howson 2015)	Students believe (HEPI survey findings) that this would be a good indicator	Cliff edge of 2.i/2.ii for employment, differing grading systems (introduction of GPA), lack of comparability in degrees between institutions. Students recognise that past performance is not always a good indicator of future performance
Output	Professional examinations pass rate (Chaffee, Darian, Gunn, Palfreyman)	External organisations that authorise delivery to their standards – comparable across the country, professional organisations already inspect and regulate	Does not cover all disciplines and all institutions
Output	Graduate placement rates (Chaffee, Gibbs, Kandiko-Howson 2013, Locke)	Student destination survey exists and shows employment rates after graduation	Gibbs – not always a good indicator for the future. Locke – employment linked to student social qualities, not necessarily to teaching quality.
Output	Graduate earnings (Forstenzer)	Can indicate progression following graduation	Is also linked to student social qualities, and affected by regulated pay in the public sector
Output	National Student Survey (Kandiko-Howson 2015, Ashwin, Gunn, Yorke et al, Holroyd and Saul) Student ratings (Berk, Barre) Student satisfaction (Chaffee, Gibbs)	Currently available Students have been the consumers of higher education, so should be well placed to judge	It is possible to game responses. It is not always a level playing field between courses. It is not always well regarded by students. There are fears of bias in responses. Student evaluations can be perceived to be superficial. Satisfaction is not the same as engagement.

Nature of metric	Description	Strengths	Weaknesses
Processes			
Process	Administrator ratings (Berk)	An American concept, related to staff engagement with the Faculty.	Not discussed in the UK.
Process	Learning gain (Gibbs, Ashwin, Schleicher, Locke)	It is possible to measure students' relationship to the subject.	Tools to measure learning gain are still in development (see below).
Process	AHELO (OECD, Schleicher, Grove)	Comparable across countries (OECD)	Not discipline related (Ashwin), and England is not taking part in the development and pilot. Doubts over the ability to compare between countries due to language differences. Risk involved in getting students to sit an externally developed test.
Process	Collegiate Learning Assessment (Massey, Grove) Proficiency Profile (Massey) Collegiate Assessment of Academic Proficiency (Massey)	Instruments in use in USA, with differences between them based on the nature of the institution that uses them.	Similar to AHELO, so there is a criticism that they are not discipline based. Low participation rates.
Process	Student engagement (Gibbs, Mandernach, Kandiko-Howson 2015, Endrizzi)	Students recognise engagement and can evaluate it, known bias can be assessed and factored out	Tools are needed to measure it, there are some which could be adapted.
Process	National Survey of Student Engagement (Massey, Mandernach)	American version – four page survey and looks beyond satisfaction to how the student engaged with the subject and with wider institution activity	Designed for use in the USA and Canada. McCormick points out that it is designed to help at institution level, not to provide an external ranking
Process	HEPI-HEA Student Academic Experience Survey (Soilemetzidis et al) UK Engagement Survey HEA (Kandiko-Howson 2015, Buckley)	UKES well regarded by students so far, does take account of what students need to put into the teaching/learning experience as well	Not full coverage yet, still under development, Academic Experience Survey not annual Engagement Survey not really intended for external ranking purposes

Nature of metric	Description	Strengths	Weaknesses
Rankings			
Ranking	Rankings (Pol, Bouchard, Wächter and Kelo, Yelland and Vale, Altbach)	Allow international comparisons, drive policy and improvement activities, international U-Multirank system allows some personalisation	Biased towards research strengths, older establishments, little information on teaching quality, reinforces existing reputations.
Qualitative			
Qualitative	Peer ratings (Berk, Thomas et al, Gast, Grove)	Good tool for improving teaching quality, essential to support metrics	Difficult to turn into a metric, sample sizes within institutions, different views on how to teach between institutions
Qualitative	Self-evaluation (Berk) Videos (Berk)	American tools for evaluating individual teachers	Too detailed for an institution level assessment
Qualitative	Teaching scholarship (Berk)	Does emphasise thinking about quality of teaching	Only part of research output and may be programme specific
Qualitative	Student interviews (Berk)	American tool for evaluating individual teachers	Expensive and challenging to turn into an institution level assessment
Qualitative	Academic Audit (Massey)	Gives depth to quantitative metrics	Generally done on a five or ten year cycle, evidence that can be influenced by institution reputation rather than what is observed.
Qualitative	QAA ratings (HEPI, Gibbs)	Already exist for institutions in England	Students are not particularly interested in these as a metric, more focussed on quality processes than teaching.

Annex 2 - Focus Group Topic Guide

Introduction / warm up

- Inform students about the purpose/nature of the research project and give information about the team.
- Explain who is funding the project and why it is being carried.
- Explain that the purpose of this focus group is to hear students' honest, open views on teaching quality.
- Briefly explain that by 'teaching quality' we mean the whole aspect of teaching – i.e. course and curriculum design, teaching intensity, stretch/challenge and support for retention and progression for students without a family tradition of going into HE]
- Summarise the themes and issues that are going to be discussed
- Give assurances of anonymity and confidentiality. Make it clear that whilst you will be making notes, responses will only be reported in aggregate and no individual or individual organisation will be named.
- Explain the rights of individuals in relation to participation on the project.
- Give assurances concerning the retention, storage and use of information.
- Make explicit the procedures in case of complaint.
- Quickly review the students attending - what subject are you studying? What year are you in?

Factors students considered when deciding which universities to apply to

1. What factors did you consider when deciding which universities to apply to?
2. To what extent did you consider teaching quality – [by which we mean repeat as above] as a factor when you were applying to university? How important was this compared to other factors?
3. What, if any, sources of information about teaching quality did you access?
4. What information was provided, or you were able to access, from HEIs themselves and did you find this useful?
    - How easy did you find it to obtain reliable information on the *teaching quality* at the institution and course you were interested in?
    - How easy did you find it to obtain reliable information on *employment outcomes* at the institution and course you were interested in?
    - How easy did you find it to obtain reliable information on the *A-level / qualifications* needed to best secure a place at the institution and course you were interested in?
5. How has your experiences of teaching quality differed from your expectations when you applied? Has the quality of teaching met your expectations so far? If not, why? If exceeded or fallen short, in what way? How does the quality compare to your pre-HE teaching experiences?

How TEF may affect the students' behaviour and decisions

6. There is a proposal to introduce a new set of incentives whereby experts will provide a rating of the quality of teaching at different universities based on student satisfaction, student retention rates and graduate job prospects. This is to ensure that HEIs are measured according to the outcomes and progress their students make. How useful do you think this information on teaching quality would have been when you were applying to university?
7. In the future, it is possible that this rating will be available at a discipline/subject level information as well as institutional level. How useful do you think this would have been when you were applying to university?
8. To what extent do you think the TEF is likely to be useful to people applying to universities in the future?

Current views on teaching quality within HE

9. To what extent do you feel that you have a sufficient number of contact hours with the appropriate level of academic staff to support you in your studies?
10. How suitable do you feel lecture and seminar class/ tutor group sizes are for the needs of your course?
11. To what extent do you feel academically stretched and challenged?
12. To what extent do you feel engaged in your learning and motivated to pursue self-study?
13. Do you feel that your course is helping you develop the knowledge and skills you need for the future? Do you feel that you are receiving adequate careers guidance?
14. To what extent do you receive feedback in a timely and helpful manner? What (if any) changes could be made to the way you receive feedback to make it more meaningful/useable?
15. How well do you feel the teaching on your course supports your wider skills and prepares you to enter employment following your studies? What more could your tutors/lecturers provide to support this?
16. Have you ever considered dropping out? What support have you received to help you stay on?
17. Overall, how satisfied are you with the quality of your course? Do you feel like you are getting value for money for your course? Why/ why not?