The Teaching Excellence Framework: Would you tell me, please, which way I ought to go from here?

Abstract
The government's Green Paper, 'Fulfilling our potential: Teaching Excellence, Social Mobility and Student Choice', presents both significant challenges and opportunities for universities. Whilst the quantitative element of the proposed Teaching Excellence Framework *(TEF), underpinned by* Big Data*, offers the tantalizing opportunity to measure, collate and compare data efficiently between all universities, it also presents a number of significant risks in the form of* institutional decision-making*, the skewing of behaviours, and potential* homogenisation of provision *across the sector. As such, this paper asserts that it is vital that the qualitative element of the TEF is implemented diligently and effectively so as to encourage universities to develop clear internal and external narratives that clearly outline the importance of education to their respective institutional missions. The authors also assert that there is an inherent danger in the prioritisation of Big Data within any framework, no matter how tempting it may prove to be in terms of uniformity of application across the sector or cost-effectiveness. Whilst such data may prove informative, it requires context in order to be truly useful. It is proposed here that the opportunity exists for universities to articulate their unique* educational missions *along with the associated* inter-relationships *between education, research, business and the professions, as well as the profile of their student bodies, so as to highlight the* 'value added' *gained by their graduates and the emerging* parity of esteem *between* education *and* research *within* Higher Education*. However, the task ahead for both the government and universities is far from straightforward.*

Dr Dan Berger, University of Hertfordshire
Professor Charles Wild, University of Hertfordshire

Introduction
The role of universities within society has been the subject of constant discussion and conjecture amongst politicians, the public, as well as within the Higher Education (HE) sector itself (Barnett 2010). However, this issue has come ever more to the forefront of people's minds in recent times due to the comprehensive spending review (CSR), related concerns regarding student fees and public debt, and governmental demands for the increased accountability of Universities in terms of student satisfaction and perceived 'value for money'.

The Research Excellence Framework (REF), which replaced the Research Assessment Exercise (RAE), is accepted generally as a reasonably effective means by which universities, and their researchers, may be assessed periodically and subsequently based on this assessment, allocated Quality Research (QR) funding in recognition, as well as to reinforce, research excellence in the HE sector. However, the results of any such exercise are subject to interpretation in a variety of ways and, inevitably, it has led to claims that the 'education' or 'teaching' function of universities has become the poor relation to the sector's research agenda (Cashmore et al 2013; Locke 2014), accompanied by claims that academics would prefer to conduct research than teach their students (Young 2006; Scott, 2015).

There have been a number of initiatives which have sought to address this perceived imbalance, and to regulate, monitor as well as to promote the educational function of universities. Amongst these have been Teaching Quality Assessments, Centres of Excellence in Teaching and Learning (CETLs), the National Student Survey (NSS), which is now enjoying its tenth anniversary, and the most recent initiative in the form of the Green Paper entitled 'Fulfilling our potential: Teaching Excellence, Social

Mobility and Student Choice' and referred to as the Teaching Excellence Framework (TEF) from the Minister of State for Universities and Science, the Right Honorable Jo Johnson MP.

In many respects, it is hard to challenge the principle that universities should be assessed in the same way for their teaching function as they are for their research activities and outputs (Geuna, 2001: 607). Indeed, the perceived success of incorporating 'impact' into the REF offers a seductive promise of similar connections that may be made with teaching in the form of progression, achievement and employability statistics. After all, this is the era of analytics and Big Data, why should it not be used as part of a new system if it is readily available across the HE sector?

The authors do not, in principle, disagree with this direction of travel. After all, a student should leave university having benefitted from the experience with a qualification that is not only recognised as being of value by employers but also of relevance to the current and future market in which graduates will find themselves. However, as with any new initiative, time and thought should not only be given to the precise objective of this exercise, but also to the potential pitfalls that may result from making a less than perfect choice. In this respect, the consultation surrounding the current green paper may be likened to the following passage from Lewis Carrol's 'Alice's Adventures in Wonderland (1865: 6.45-50):

> "Would you tell me, please, which way I ought to go from here?"
> "That depends a good deal on where you want to get to," said the Cat.
> "I don't much care where – "said Alice.
> "Then it doesn't matter which way you go," said the Cat.
> " – so long as I get somewhere," Alice added as an explanation.
> "Oh, you're sure to do that," said the Cat, "if you only walk long enough."

To take a rather simple analogy, when it comes to the educational function of universities or, more specifically, their 'teaching' of students, are we in fact comparing apples and oranges or merely green apples with red ones? The reality within England and Wales is that students are currently faced with a higher education landscape littered with coalitions and groups of universities that include everything from the research-intensive Russell Group, through to the University Alliance, the Million+ Group, and the GuildHE. In addition to this, there is an ever growing range of individual university mission statements that seek to differentiate the specific experience on offer at one HE institution from another. This has inevitably led to a polarisation of views, stereotyping of behaviours and offerings, as well as a growing concern that certain institutions favour either research or teaching at the expense of the other (Robertson & Bond: 2001). If we are to believe many of the mass media's headlines in recent years, universities are as different to one another as apples and oranges.

Whilst the notion of higher education may be traced back to the Middle Ages, the first universities emerged towards the end of the 11th Century in Paris and Bologna in response to the evolving needs of society for professional training (Hashimshony & Haina, 2006). Over time, evolving and occasionally conflicting societal demands have led both academics and politicians to question whether the role of universities needs to be redefined, in response to contemporary society. In this regard, Barnett (2010) highlights the fact that the original education-centred purpose of universities has long since evolved into a multi-faceted mission that includes research, consultancy and knowledge transfer activities, with each university possessing a variety of shared characteristics. As such, whilst Barnett identifies different types of university (metaphysical, scientific, entrepreneurial, and bureaucratic) that may be used to characterise HE institutions over time, not one single type can satisfactorily characterise any one single university; they simply highlight the dominant strains that exist within the HE sector. As such, universities should be regarded as amalgams of these types, 'within a single university can be found the research university, the entrepreneurial university, the bureaucratic university and now the corporate university' Barnett (2010: 50). Therefore, returning to our earlier analogy, whilst we are not comparing apples with oranges, the question may be posed as to whether, in fact, we are comparing green apples with red ones?

Adopting a slightly different perspective, Delanty (2001) observes that the university is the only societal institution where four inter-connected activities take place; education, research, professional training and intellectual criticism. However, as Hashimshony & Haina (2006:8) note 'the relative importance of each activity must be re-evaluated' in response to both the changing demands that have been placed on universities and the emerging challenges these institutions face. For example, the recent comprehensive spending review highlights the need for universities to be 'more responsive to the demands of [the] market, recognise the need to change their ways…and become more efficient' (Jarvis, 2000: 52). Similarly, advances in technology over the past decade have had a significant impact on modes of delivery, whether in terms of the flipped classroom model (Berger & Wild 2016) or more traditional concepts of distance learning. As Kumar (1997: 31) highlights, a key function of universities to date has been to create a 'sense of community' and, through that, to facilitate the cross-fertilisation of ideas amongst the student body. Therefore, a major challenge for those universities that have chosen to utilise technological advances is how to maintain such a community going forward.

Therefore, the resultant balance achieved between Delanty's four activities, coupled with Barnett's evolving future types of university (e.g. liquid, authentic), will have a significant impact on our understanding of their role within society, whether at a sector-wide or individual institution level. Nevertheless, universities will remain the primary centre of higher education and as a repository of accumulated knowledge and wisdom (Bell 1973); the dual objectives of discovery and dissemination.

Whilst some institutions may seek to invest significant time and effort growing, and maintaining, a reputation for strength in research, every university should excel in their educational functions, irrespective of whether they happen to be a part of the Russell Group, University Alliance, or Million+ Group. In this regard, the authors assert that the government's proposed Teaching Excellence Framework is designed precisely to achieve this result by ensuring that education is a priority amongst all universities, irrespective of whether they are scientific, bureaucratic or entrepreneurial. Indeed, this lies at the heart of our understanding of the driving force behind the Green Paper. The framework seeks to extend further than that of measuring individual teaching practice within universities and is based on the premise that the more rounded concept of 'education' is a more appropriate concept to use as the basis of measurement than that of 'learning and teaching'.

In this regard, Fung & Gordon (2016: 5) note that there needs to be a widespread acceptance of the term 'education' amongst both academics and politicians so as to capture more accurately a university's collective educational mission which, as noted above, includes the inter-relationships between research, business and the professions with education, rather than focusing solely on individual 'teaching' practice. In many respects, this approach starts to mirror that of the REF, which seeks to look not only at research output but also its inter-relationships.

However, the era of analytics and Big Data offers the tantalizing opportunity to measure, collate and compare data between universities relating to student retention, progression, achievement on graduation and employability statistics. In addition, there is data relating to the ethnic and gender profiles of student cohorts, information relating to widening participation, statistics relating to student satisfaction at module and programme level, as well as data relating to the student views of individual lecturers. In a perfect world, there is data that covers almost every aspect of a student's time at university; data which could easily be regarded as capturing the educational mission of a university. The reality though is that we do not exist within a perfect world. Whilst information, data and statistics are useful, they are only truly helpful when they form part of a more detailed picture.

To date, the proposal is for the quantitative aspect of the framework to focus on a university's Quality Assurance Agency (QAA) results, annual NSS scores, employability rates and programme retention rates. Whilst such data sets undoubtedly capture aspects of a university's educational mission, the authors assert that they are also more aligned to the concepts of 'value for money' and 'what employers want', as contained in the Green Paper itself, than the notion of 'teaching excellence' as highlighted in the framework's title. Nevertheless, the adoption of such a standardised system ensures that when it comes to measuring education within the universities, we are in fact starting from a

baseline of comparing green apples with other green apples. However, as with any new system, this inevitably raises a number of questions that need to be clarified prior to implementation, including the very obvious one of 'what, for the purposes of our comparison, is green?' and, once that is determined, 'what level of divergence from green is considered acceptable, and by whom?'

Consequently, whilst the quantitative element of the TEF will certainly play an important role going forward, it should neither form the sole basis of the government's strategy nor dominate a university's strategic response to the process in terms of silo strategies targeted at each data set, as per Barnett's bureaucratic or corporate university models. It is vital that the qualitative element of the TEF is implemented effectively so as to provide a holistic approach which emphasises each university's individual mission, the unique inter-relationships between education, research, business and the professions within institutions, as well as the profile of a university's student body in terms of gender, ethnicity and socio-economic background, so as to highlight the 'value added' gained by graduates of that particular university (Biggs 1993).

In this respect, we are Alice at the crossroads. Should we pursue a course that places complete faith in the power of Big Data, which is dominated by the quantitative aspects of the proposed TEF, and which ultimately produces rankings between universities that are readily accessible to the mass media? Or do we ensure that the system's qualitative element is used effectively so as to provide a useful, and balanced, insight into the underlying culture and ethos of educational provision offered by each individual HE provider, and accept that whilst there are significant overlaps between universities Big Data can, in practice, only go so far in terms of reinforcing a picture, as opposed to driving it? The authors suggest that we must ensure that the system is able to appreciate that whilst all universities may appear as green apples according to the quantitative data element, not all green apples perform the same function; some are for eating whilst others are for cooking; something which only the qualitative element can address. To continue the quote from Lewis Carrol's 'Alice's Adventures in Wonderland (1865; 6.45-50):

> `What sort of people live about here?'
> `In THAT direction,' the Cat said, waving its right paw round, `lives a Hatter: and in THAT direction,' waving the other paw, `lives a March Hare. Visit either you like: they're both mad.'
> `But I don't want to go among mad people,' Alice remarked.
> `Oh, you can't help that,' said the Cat: `we're all mad here. I'm mad. You're mad.'
> `How do you know I'm mad?' said Alice.
> `You must be,' said the Cat, `or you wouldn't have come here.'

We are faced with two extremely challenging, if not slightly mad, routes to measuring the educational missions within universities. The former quantitative route will be popular, if not with the HE sector, certainly with central governmental departments and the mass media. It is the path of least resistance due to the fact that these data sets already exist and have been widely, if somewhat reluctantly, accepted as an integral part of the annual reporting cycle for universities. The potential danger with Big Data is that it simply tells you whether one type of green apple, most probably the eating apple, is more attuned to the exercise than another. If we return to Barnett's (2010) university models, the scientific university (i.e. a research-intensive Russell Group university) has proved extremely successful under the REF. In much the same way, it may be that either a bureaucratic or corporate university could prove more adept at responding strategically to the TEF in the short term. The temptation though may be to focus on the quantitative exercise, as opposed to focusing on the far more challenging task of the overall quality of its educational mission and provision. Consequently, if Big Data is used as the sole basis for broader decision-making, reliance on data sets poses a real threat of skewing behaviours, and homogenising provision, within the HE market to the detriment of other core aspects of society that rely upon the presence of certain types of university with different educational missions.

The qualitative route is also far from straightforward. Fung & Gordon (2016: 4) highlight the challenges faced by many research-intensive universities in terms of clarifying, or even developing,

distinctive models of academic educational leadership and research-based education so as to 'create the cultural changes needed for education to have 'parity of esteem' with research'. Indeed, one of their core recommendations which, in the authors' opinion should be pursued by all types of university, is to 'develop a credible and persuasive narrative regarding the importance of education to the institutional mission' and to ensure that it is 'reiterated consistently to internal and external audiences' (Fung & Gordon 2016: 6) so as to not only effect change to internal reward and recognition processes but also to bring institutional missions relating to research and education closer together so as to encourage widespread research-based education (Brew 2006). It is also worth noting Macfarlane's (2015) observation that research-intensive universities are a relatively recent phenomenon created by a response to the RAE which 'offered a relatively easily quantifiable set of measures for reward, at least in comparison with challenges of measuring excellence in education' (Fung & Gordon 2016: 15). Once again, this highlights the potential pitfalls associated with the introduction of systems with overly dominant quantitative elements that evoke an overly narrow response from the bureaucratic or corporate aspects of a university and the need to keep TEF's qualitative element at the forefront of discussion.

Going back to our earlier analogy, whilst all universities may appear as green apples for the purposes of comparing the quantitative element of their educational functions, not all green apples perform the same function; some are for eating whilst others are for cooking and it is the qualitative element of TEF, underpinned by clear institutional narratives, that will provide the opportunity to clarify this. It also provides a unique opportunity for universities of every type and varying mission statement to reflect not only on their collective educational missions but also on the inter-relationship that could exist more effectively between education and research (Robertson & Bond: 2001) but which, over time, have become separate strands within HE institutions due to funding and policy considerations.

Consequently, it would prove to be 'madness' to pursue a particular path without giving due consideration to the broader implications of the eventual system to be adopted. In addition, it should be borne in mind that there is an inherent risk associated with any new system, such as with the Teaching Quality Assessment (TQA), in that 'the management of quality [eventually takes] over from the delivery of quality' (Underwood 2015: 4).

With this in mind, the remainder of this article will seek to outline past experiences of quality assurance, and to suggest lessons that need to be borne in mind for the current consultation exercise. If one considers the Green Paper, key terms such as 'value for money' or 'what employers want' appear 27 and 35 times respectively, whilst the terms 'teachers' and 'academic' both feature twice, Whilst a great deal is understandably reliant on both choice of language and context, it is interesting to note that the former two terms mirror those used by Harvey and Green (1993) in their discussion of conceptualisations of quality within education, as well as in 'Dimensions of Quality' (Gibbs 2010). The authors assert that we should seek to implement a framework that does not give in to the temptation of Big Data but rather seeks to take account of the complexities outlined in Biggs' '3P' model (1993), so as to gain a more accurate picture of an institution's educational provision.

The Teaching Quality Assessment and Quality Assurance Agency

Quality assurance within HE as a distinct area of activity, may be traced back to the late 1980s with the publication of the Reynolds Reports (1986) on the procedures within universities designed to monitor and guarantee standards in the content of their degree programmes and the quality of their teaching. These reports highlighted a number of areas and related recommendations which, today, are a central part of any HE institution's day-to-day activity, and included external examiners; external involvement in the maintenance and monitoring of academic standards; appeals procedures; and internal procedures whereby academic staff' were monitored in terms of the effectiveness of their teaching. However, these recommendations were voluntary and not embraced as enthusiastically as they could have been by those involved. As such, in place of a mutually agreed voluntary code of practice, universities were faced in the early 1990s with institutional level Academic Audits and, shortly afterwards, subject level Quality Assessment, later renamed Subject Review.

A key point to be noted for the current consultation exercise is that when the Subject Review process commenced, it included auditors undertaking 'the observation of samples of all forms of teaching' without prior notification of who would be chosen. The impact of this was that universities sought to engage everybody in a department with the aim of ensuring that they reflected on how they did their teaching and reviewed / refreshed their teaching materials. One of the fundamental misconceptions of the new TEF is that it does not actually seek to measure teaching practice or its excellence as per the Subject Review process. Instead, it is focused on the more holistic educational mission of a university and some of the major associated outputs of that process such as progression, employability, and student satisfaction. If the government is not careful in its implementation of the new framework, many academics could feel that they are at least one level removed from such outputs, many of which are produced by central university administrators and external governmental officials, as opposed to university lecturers themselves.

By contrast, despite the bureaucracy associated with the REF is it based on the peer-review of an individual's research output within a university by panels of researchers involved in the very process they are seeking to grade (Times Higher 2015). It retains a very personal element for the academic involved which, if the TEF is to prove successful must be borne in mind by both the government as well as individual universities so as to ensure that the educational community is engaged from the outset. As Graham (2016: 6) observes, 'the primary focus to date has been on the quality of research; motivating, measuring and rewarding research excellence. The spotlight is now turning to teaching quality.'

This relates to another aspect of the TQA and the observation that 'the management of quality [eventually] took over from the delivery of quality' (Underwood 2015: 4). As Sir David Watson later noted, 'the [quality assurance] war itself distracted us from improving teaching as much as we could have done' (Watson 2006: 6). In other words, disagreement as to the shape and form of quality assurance within HE meant that the core objective of the exercise – the interests of students – was eventually sacrificed to the management of the process, resulting in situations whereby an institutions 'quality assurance could be good even if teaching quality wasn't' (2015: 4).

One of the reasons that this aspect was eventually discontinued in favour of a greater focus on quality assurance was the fact that it was never truly accepted. Universities are inevitably far more complex and dynamic places that schools or colleges, meaning that Ofsted style inspections needed to be far more sensitive to the differing learning and teaching strategies adopted by one institution to another. Equally, as Graham (2016: 6) notes the metrics used for evaluating the educational contribution of academics are 'widely understood to be poor indicators of teaching quality and hold the confidence of neither promotion candidates not their university managers.'

Indeed, if one considers the Quality Assurance Agency (QAA) which was intended to combine the two streams of quality assurance - Academic Audits and Subject Review, it too has faced a number of challenges over the past twenty years. Initial attempts by the QAA to establish an Academic Infrastructure focused on the provision of assurance about standards and quality (consisting of an HE qualifications framework, Code of Practice and Subject Benchmark Statements) (Dearing Report: 1997) were eventually replaced by a revised Quality Code between 2011 and 2013.

Similarly, early attempts to introduce a sector wide process of Academic Review (comprised of elements of both Subject Review and Academic Audit), faced objections from universities as to the administrative burden that such a process would impose, resulting in a modified approach, known as Institutional Audit, being devised. This too was quickly modified so as to ultimately remove the 'discipline audit trail' (DAT) element and to provide space within the audit process for panels to explore a broader range of topics and themes. As such, the original intention of the DAT which had been to provide selective subject-based enquiries that enabled a phased reduction of the subject focus of QAA reviews, but which would still retain some form of Subject Review, was ultimately lost within a decade of the QAA's introduction.

Whilst a new method of Institutional Review has subsequently been developed, questions persist on the one hand as to whether there should be a further shift in focus towards a risk based approach (so as to alleviate the administrative burden on HE institutions further), or whether established universities should be the subject to a lighter touch than further education colleges or new private providers. On the other hand, there are still questions as to the effectiveness of the QAA's quality assurance system, with one of the most notable criticisms coming from the chairman of the House of Commons' Select Committee on Universities who condemned the Agency as 'a toothless old dog' and declared that the British degree classification system had 'descended into farce' (Adams: 2008). Yet the QAA has always maintained that 'the quality of the education offered by UK institutions is its strength and the basis of its strong reputation' (BIS Committee, 12th Report 2010-12).

Throughout this debate, it is worth remembering that the 'QAA (unlike Ofsted) does not directly assess quality: it assesses how autonomous universities assess the quality of what they themselves do' (McGee 2014). This is reflected in the Report issued by the Science and Technology Committee of the House of Lords (2012), which noted 'individual institutions are autonomous organisations and, as such, have primary responsibility for academic standards and quality. QAA provides a means of external assurance to this process' (Science and Technology Committee: 2012 Chapter 5, para 109). However, the Committee ultimately concluded that the QAA was still not fit for purpose with its Report centred on the fact that reviews are based on a 'threshold level' of standards that 'allowed no assessment of quality above that threshold' (Science and Technology Committee: 2012 Chapter 5, para 124) and that more needed to be done to improve quality (Science and Technology Committee: 2012 Chapter 5, para 125). In other words, whilst the current system enables us to identify that we have an apple, which is claimed to be green, satisfies the threshold for 'greenness' but for which confirmation beyond this point is not possible.

The Report makes interesting reading, especially at a time when we are considering the possible routes that could be followed when implementing the new TEF. The question arises as to whether a more rigorous, centralised approach should be adopted and, if so, the potential impact that this could have on HE institutions and the inherent diversity of the sector. For example, one of the positives always afforded to the QAA is that 'it [currently] respects the missions of different institutions, in a way that the research excellence framework, league tables and newspapers do not' (McGee 2014). The question is whether a more rigorous approach to quality would be desirable, or indeed possible? Would a system that seeks to determine 'what green is' and 'what level of divergence from green is considered acceptable' impact negatively on the diversity of the HE sector or the ultimate goal of all involved – the education of students? The starting point is whether we can define quality within the context of higher education.

Measurements / Thresholds of quality

The QAA seeks to make a distinction between standards and quality. It states that a standard represents the level of achievement that a student needs to reach in order to gain an academic award. By contrast, it defines quality as a way of describing 'how well the learning opportunities available to students are managed so as to help them to achieve their award. It is about making sure that appropriate and effective teaching, support, assessment and learning opportunities are provided for students' (Science and Technology Committee: 2012 Chapter 5, para 112).

There have been a number of attempts to define the concept and nature of 'quality' within the area of education (Cheng and Tam, 1997). However, one of the most commonly utilised is that provided by Harvey and Green (1993) who suggest that quality should be regarded as a relative concept, not only in terms of meaning different things to different people but also in relation to processes and outcomes. Gibbs (2010: 11) goes on to note that 'what matters is whether one educational context has more or less quality than another, not whether it meets an absolute threshold standard so that it can be seen to be of adequate quality, nor whether it is reaches a high threshold and can be viewed as outstanding and of exceptional quality, nor whether a context is perfect, with no defects.'

Harvey and Green (1993) go on to group conceptualisations of quality into five categories; exception, perfection, fitness for purpose, value for money, and as transformative. As noted earlier, the latter two categories are of particular interest in the context of the TEF. The notion of 'value for money' is something which lies at the heart of the Green Paper and, many would suggest, appears to be definition of quality driving the government's agenda in this area. The suggestion of quality as being transformative or, put more simply, the 'educational gain' of students during the course of their studies (as opposed to their ultimate educational performance) is also of particular interest for a number of reasons.

First of all, Cambridge University's recent response to the TEF consultation exercise that 'the 'long-reach' aim of universities is to help students grow into thoughtful and critical citizens' as opposed to merely enabling them to become 'earners and consumers', mirrors the authors' stance of the significance and importance of developing students' critical reasoning skills (Berger & Wild 2016). Therefore, any measurement of quality must be linked to this aspect of a university education. The other key term used throughout the Green Paper is 'what employers' want' and, again, may be linked with Cambridge's response. The aim of education is to instill more than just subject-specific knowledge or technical ability; it is to develop a student's 'problem-solving, creativity, teamwork and critical thinking' (Morgan 2016), which are precisely the skills and attributes that employers value. Finally, the notion of quality as being measured in terms of transformative impact is of particular relevance when considering the validity of student judgments of the quality of teaching they have received whilst at University, particularly in relation to another of the government's current measures of quality; the NSS. Student perceptions of what they would like to receive from their lecturers may not always correspond with 'what is known to work in terms of educational effectiveness' (Gibbs 2010: 11). Equally, student interpretations of the questions being posed not only vary between those of academics and students but also amongst students themselves due to the fact that 'survey questions are too vague to elicit meaningful answers' (Agrawal, Buckley-Irvine & Clewlow 2014)

In addition to this, the notion of 'teaching excellence' remains rather elusive as many of the metrics used for evaluating the educational contribution of academics are, as noted earlier, widely perceived to be poor indicators with little confidence placed in them by either academics or their managers. In this regard, Graham (2016: 12) suggests that any education-focused framework could learn valuable lessons from the success of research achievement measures, which are not only 'well understood and recognised' but have the inherent flexibility to recognise and reward 'different types of research contribution…to industrial impact or to prestigious publications'. As such, education achievement measures should offer a similar degree of transparency and flexibility so as to encourage greater engagement from the academic community. Graham outlines a template that provides definitions and criteria for 'effective teachers' as a baseline through to 'skilled and collegial teachers', two potential routes of 'scholarly teachers' of 'institutional leaders' - 'one focused on impact on the educational environment and one focused on impact on educational knowledge' (2016: 14) – leading to the final category of 'national and global leaders'.

Whilst this provides a useful starting point for identifying 'teaching excellence' at an individual level, the challenge remains as to how a robust, yet flexible and cost effective, system for gathering evidence to support the educational achievement of academics may be implemented. Amongst the forms suggested by Graham (2016) are indirect measures such as rates of progression and attrition, student feedback and post-graduation employment rates; the precise data sets outlined in the government's consultation document as forming the framework's quantitative element. Indeed, at a recent debate organised by the Higher Education Policy Institute and the Higher Education Academy, Lord Willetts said that the proposed TEF metrics of student retention, graduate employment and NSS results go 'to the heart of the kind of issues that students care about' (Havergal 2016), going on to qualify this statement by noting that it's important that such metrics are considered in their proper context. Indeed, it is this qualifying statement that is at the heart of any consideration of the Green Paper – context and complexity. In this regard, Graham suggests additional sources of evidence which should be included, including peer evaluation, self-assessment, and the professional activities of academics are also included along with more direct measures of student learning such as 'learning

gain over a period of time' (Graham 2016: 29). It is these latter sources of evidence which could usefully form part of the TEF's qualitative element.

What is clear is that education is an extremely complex process and, as such, it is necessary to seek to understand the dimensions of quality within education along with the varied contexts and interactions of these elements. In this regard, the '3P' model (Biggs 1993) treats education as a complex system in which the three dimensions of quality (presage, process and product) interact with one another. Biggs' model is seen traditionally as the most effective means of starting to analyse the education environment. Within the '3P' model:

- Presage defines the HE institutional context that exists before a student commences his/her learning, such as resources, the quality of the students or academic staff. These variable are seen to frame, enable or constrain the form education takes within an institution;
- Process defines variables that affect the student's learning experience, such as class size, class contact hours, the extent of feedback from academic staff; and
- Product variables relate to the outcomes of the educational process and include student performance, retention and progression rates as well as employability. Occasionally, these may also include the notion of educational gain (G.Gibbs 2010: 11).

What is readily apparent from this list is the fact that each and every category is already the subject of data collection by universities or government agencies, including those highlighted in the Green Paper; student retention, graduate employment and the NSS. Data also exists, and is collated annually, for Key Information Set (KIS) which outline information relating to class contact hours, student achievement, average entry tariffs, professional accreditation as well as graduate employability rates. As noted earlier, the era of analytics and Big Data offers the tantalizing opportunity to measure, collate and compare data between universities for all of these areas.

However, as Lord Willetts observes, whilst statistics are informative, they are only truly useful when they form part of a more detailed, complex and contextualised picture. In reality there is considerable overlap and interaction between the three Ps outlined in Bigg's model which contributes to an extremely complex system. Indeed, there have been attempts to subdivide these areas into more detailed categories (Finnie and Usher, 2005; Usher and Savino, 2006) but the reality is that we are far from being able to model the arena of education effectively or usefully. Indeed, the Science and Technology Committee's Report ultimately notes that, in their opinion, the quality of a degree is dependent on all three factors 'to ensure that high quality graduates leave Higher Education with the right skills and knowledge to prepare them for work' (2012, Chapter 5, para 113).

"I want to get to Avalon" said Alice

In principle, the proposed Teaching Excellence Framework is a commendable idea, insofar as there is a genuine need to ensure that all students from a wide range of backgrounds, attending a wide range of HE institutions receive the best quality education to which they are entitled. However, any consideration of universities, education, teaching and learning or indeed the notion of excellence, are far from straightforward.

It is clear that the original education-centred purpose of universities has evolved over time in response to the changing needs and demands of society so as to form a multi-faceted mission that includes research, consultancy and knowledge transfer activities. Alongside this, academics such as Barnett (2010) suggest that whilst each university possesses a variety of shared characteristics, the way in which these institutions interact with the market goes beyond their individual mission statements and lies at the very heart of these organisations in terms of whether they may be characterised as scientific, bureaucratic or entrepreneurial. Consequently, whilst the underlying multi-faceted mission remains a constant, the way in which it is balanced and implemented varies considerably between institutions. Consequently, it has led many to question whether, when comparing a Russel Group university with an Alliance university, we are in fact seeking to compare apples with oranges.

The authors suggest that the TEF has sought to provide a benchmark position that it applicable to all universities, regardless of their type or specific mission. This takes the form of a quantitative element that draws on existing data sets that capture indirect measures of student learning in the form of student progression and retention, student feedback results and graduate employability. It also seeks to provide a framework in which we are in fact comparing green apples with other green apples. However, the authors assert that there is an inherent danger in the prioritisation of Big Data within any framework, no matter how tempting it may prove to be in terms of uniformity of application across the sector or cost-effectiveness. Whilst such data may prove informative, it requires context in order to be truly useful.

It is at this point that Fung & Gordon's (2016) suggestion that there needs to be a widespread acceptance of the term 'education', rather than continuing to focus on individual 'teaching' practice, so as to capture more accurately a University's collective educational mission and to understand more clearly the student learning experience. If this shift in focus is undertaken, and accepted as the driving force behind the TEF, then there are a number of consequences for the qualitative element of the proposed framework.

First of all, by focusing on an institution's educational mission, it includes the inter-relationships between research, business and the professions with education. In other words, it accepts the multi-faceted mission of the modern university and provides an opportunity for institutions to clarify the implications for a student's learning experience. Secondly, it provides a unique opportunity for universities to clarify, or even develop, distinctive models of educational leadership and research-based education so as to create an internal environment where education has 'parity of esteem' with research. Linked with this, is the need to bring institutional missions relating to research and education closer together so as to encourage widespread research-based education (Brew 2006). Thirdly, the adoption of a holistic approach should mitigate the potential risks associated with Big Data in terms of internal university decision-making leading to the skewing of behaviours and homogenization of provision in an effort to excel at specific data gathering exercises, by focusing on the entirety of a university's education mission rather than a limited number of outputs.

In many respects, this links with the authors' final point. It is vital for all universities, irrespective of type or mission, to develop a clear, accessible narrative that clearly outlines the importance of education to their institutional missions. Such narratives will not only form the basis of institutional responses to TEF's qualitative element, but provide the opportunity to clarify each university's individual mission, the unique inter-relationships between education, research, business and the professions within different institutions, as well as the profile of a university's student body in terms of gender, ethnicity and socio-economic background, so as to highlight the 'value added' gained by graduates of that particular university. It is this aspect of the TEF that will, if implemented carefully, provide the opportunity to acknowledge that not all green apples are the same and that, in fact, society needs variety of provision. Our destination must be Avalon, a land of apples but also a land of different types of apple whereby differences are acknowledged and valued.

Bibliography

Adams.S, (2008), 'University degree system 'is a farce', The Telegraph: News: Telegraph Media Group Limited, retrieved 3 March 2016

Agrawal.A, Buckley-Irvine.N & Clewlow.E, (2014), 'Is the National Student Survey fit for purpose?' The Guardian, retrieved 3 March 2016

Barnett.R, (2010), 'Being a University', London and New York: Routledge

Bell.B, (1973), 'The Coming of Post-Industrial Society: A Venture in Social Forecasting', New York: Basic Books

Berger.D & Wild.C, (2016), 'Refining the traditional flipped-classroom model to optimise student performance on undergraduate degree programmes', Journal of Commonwealth Law & Legal Education, forthcoming

Brew.A, (2006), 'Research and teaching: Beyond the divide', London: Palgrave-MacMillan

BIS Committee, 12th Report (2010-12): Government reform of Higher Education (HC 885)

Cashmore.A, Cane.C & Cane.R, (2013), 'Rebalancing promotion in the HE sector: Is teaching excellence being rewarded?' York: Higher Education Academy

Cheng, Y .C & Tam, W .M, (1997) 'Multi-models of quality in education', Quality Assurance in Education 5 (1), pp22–31

Dearing Report, (1997), 'Higher Education in the Learning Society', London: Her Magesty's Stationery Office

Delanty. G, (2001), 'Challenging Knowledge: The University in the Knowledge Society', Buckingham: Open University Press.

Etzkowitz.H, Webster.A, Gebhardt.C, Terra.B, (2000), 'The future of the University and the University of the future: Evolution of Ivory Tower to Entreprenurial Paradigm', Research Policy 29(2), pp313-330

Finnie, R & Usher, A, (2005), 'Measuring the Quality of Post-secondary Education: Concepts, Current Practices and a Strategic Plan', Kingston, Canadian Policy Research Networks.

Fung.D & Gordon.C, (2016), 'Rewarding educators and education leaders in research-intensive universities', Higher Education Academy

Geuna.A, (2001), 'The Changing Rationale for European University Research Funding: Are There Negative Unintended Consequences?', Journal of Economic Issues 35(3), pp 607-632

Graham.R, (2016), 'Does teaching advance your academic career? Interim Report on the development of a template for evaluation teaching achievement', Royal Academy of Engineering

Higher Education Directorate, Department for Business, Innovation & Skills (November 2015), 'Fulfilling our potential: Teaching Excellence, Social Mobility and Student Choice', available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474266/BIS-15-623-fulfilling-our-potential-teaching-excellence-social-mobility-and-student-choice-accessible.pdf

Harvey, L & Green, D, (1993), 'Defining quality', Assessment and Evaluation in Higher Education 18 (1), pp9–34

Hashimshony, R. & Haina, J. (2006), 'Designing the University of the Future', Planning for Higher Education, 34(2) pp5–19

Havergal.C, (2016), 'No excuses on teaching quality with £9k fees, says Willetts', Times Higher Education, retrieved 3 March 2016

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

Jarvis.P, (2000), 'The Changing University: Meeting a Need and Needing to Change', Higher Education Quarterly 54 (1) pp 43–67.

Kumar.K, (1997), 'The Need for Place', in eds.Smith.A & Webster.F, 'The Postmodern University? Contested Visions of Higher Education in Society', Buckingham: Open University Press, pp 27-35

Locke.W, (2014), 'Shifting academic careers: implications for enhancing professionalism in teaching and supporting learning', York: Higher Education Academy

Macfarlane.B, (2015), 'Academic practice, identity and careers', Society of Research in Higher Education, Special Issue 1965-2015.

McGhee.P, (2014), 'What has QAA ever done for UK universities? A lot, actually', The Guardian, retrieved 3 March 2016

McLeod.D, (2004), '£300m set aside for teaching excellence centres', The Guardian, retrieved 3 March 2016

Morgan.A, (2016), 'University of Cambridge: Green Paper plans risk 'considerable damage'', Times Higher Education, retrieved 3 March 2016

Robertson.J & Bond.C, (2001), 'Experiences of the relation between teaching and research: What do academics value?', Higher Education Research and Development 20(1) pp5-19

Science and Technology Committee, "Second Report on Higher Education in Science, Technology, Engineering and Mathematics (STEM) subjects, Chapter 5". Publications.parliament.uk. Retrieved March 11, 2014.

Scott.P, (2015), 'Three reasons why the Teaching Excellence Framework won't work', The Guardian, retrieved 3 March 2016

Usher, A & Savino, M, (2006), 'A world of difference: a global survey of university league tables', Toronto, Educational Policy Institute. Available from: www.educationalpolicy.org/pdf/world-ofdifference-200602162.pdf

Underwood.S, (2015), 'Reforming quality assurance in higher education: Putting students at the centre', Centre for Analysis of Risk and Regulation

Watson.D, (2006), 'Who killed what in the quality wars?' Cheltenham: QAA

Yorke, M . (2009), 'Trends in honours degree classifications, 1994–95 to 2006–07, for England, Wales and Northern Ireland', York: Higher Education Academy.

Young.P, (2006), 'Out of balance: Lecturers' perception of differential status and rewards in relation to teaching and research', Teaching & Higher Education 11(2) pp 191-202

Response Sheet

Reviewer 1	
This is a very insightful and constructively critical article. The topic is timely and of major relevance. The argument is balanced and logical. I would suggest minor revisions before the article is published. Much of the article looks like an excellent commentary rather than an article for an academic journal. It would be important to contextualise the article much more in the academic literature including the literature on quality. References to the REF and references to the unintended consequences of the REF should also be drawn on. The article could also draw more on empirical studies on quality eg the work by John Brennan and colleagues to bolster certain aspects of the arguments. The article needs careful proof reading and the removal of words such as 'don't'	All comments and suggestions for amendment have been accepted and incorporated into the text. The article now includes literature relating to quality (pp 7-8) and to the RAE / REF (pp 4-6) Proof-reading has been carefully undertaken with words such as 'don't' removed.
Reviewer 2	
Given that you are presenting an analytical critique of the directions of current policy, the analysis has a couple of weak areas which could very usefully be strengthened. On page 2 your write: *The fundamental reality is that every University should, at its core, have the dual objectives of both discovery and dissemination.* It would strengthen your argument if you cited texts which address and critique the *purpose* of the modern university – see e.g. work by Ron Barnett (such as Barnett, R. (2011) *Being a University* Routledge, Oxford and New York). To suppose that there is a 'fundamental reality' which just speaks for itself is problematic; referring for example to Barnett's notion of an 'authentic university' would add depth. Similarly, you could usefully analyse more explicitly the issue of whether the Teaching Excellence Framework is in fact about teaching, or is it (should it be?) about the quality of student education overall. You do begin to raise this, and return to it with the reference to Biggs (who interestingly doesn't really address fundamental values in his work), but don't follow up your own line of argument. You then offer a comment on 'teaching observations' which doesn't support the core argument, in part because you haven't bottomed out the relationship between teaching *per se* and student	All comments and suggestions for amendment have been accepted and incorporated into the text. The authors would like to thank the reviewer for taking the time to suggest additional elements for inclusion, which have proved extremely useful in the production of a revised version. A discussion of the literature relating to the purpose of the modern university (including Barnett) has been included (pp2-4) A discussion of the true purpose of the TEF has been undertaken and embedded throughout the article. This is most clearly seen on p3 and p8. Reference to teaching observations has been removed from the article.

education in the round (which includes student engagement and the wider student experience). In other words, the quality of classroom teaching is only one element in the whole student educational experience.	This has been addressed throughout the article and is most clearly evident on pp3-4 and p8.
You also refer, importantly, to the relationship between education and research in universities. You could make more of this by explaining further the potential synergies and dangers of splitting the two areas of focus (as is happening with the REF/TEF split). See for example the arguments in Fung and Gordon's 'Rewarding educators and education leaders in research-intensive universities' (2016): https://www.heacademy.ac.uk/sites/default/files/rewarding_educators_and_education_leaders.pdf Rewarding educators and education leaders in research ... www.heacademy.ac.uk	This has been addressed throughout. Please see pp3-4 This has been included
5 Terminology and definitions In this study we use the word 'education' rather than the phrase 'teaching and learning', and 'educator' rather than	This too has been incorporated into the discussion (p3).
Related to the core argument is that of making judgments about 'good teaching'; these are picked up in the above text and also in Ruth Graham's very recent publication, interestingly endorsed in a forward by Jo Johnson: http://www.raeng.org.uk/news/news-releases/2016/february/does-teaching-advance-your-academic-career-academy	This has been included in the discussion (pp8-9)
These or other related publications could help you bottom out the core issue of defining/problematising what is apparently meant a) by 'teaching' and b) by 'excellence' in the Green Paper, and of then explaining more fully how these terms might better be defined and understood. As you explain in the section on big data, information is by no means everything – we need to understand our values and purposes at a much more fundamental level.	This has been addressed throughout the article.
Throughout the article: your citation style is inconsistent, as is the style of your reference list. Please amend so that they consistently	This has been undertaken throughout the text

follow the correct journal style.	
At times your style is quite casual, which makes it readable but could also lack precision for readers not familiar with the material or context. For example, on page 1 you refer to 'Jo Johnson': I feel you should include his role/title, and also in your reference to Margaret Hodge, and review the article in this vein to make sure that readers not familiar with the British context or the particular nomenclature understand fully the meanings of your terms.	This has been corrected
Page 1 Para 2 - 'ill-fated': this term either needs explanation or to be changed to something less emotive.	This has been amended and supported by academic opinion
Page 6 – 'Unless the TEF includes the observation of teaching, it will never deliver what it promises.' This sits uncomfortably – there is no supporting analysis of the usefulness/limitations of teaching observations.	This has been removed and the argument revised so as to prove consistent with earlier discussion
Page 10/11 – Final two paragraphs. The last paragraph (two sentences) provides an odd ending. A summing up of the core argument is needed in a more substantial final paragraph, and for me it needs to link the threads of what being a 'good' university might mean – perhaps in opposition to the notion of 'Teaching Excellence' – and how that might be measured – if indeed it needs to be measured. The notion of 'good' needs reference to values (e.g. those of critical thinking, discovery, dialogue); the notion of measuring what is good needs reference to criteria and feasible approaches. You allude to some of these (e.g. using Biggs) but the threads get a bit lost at the end: a stronger final paragraph could make the overall narrative much more persuasive.	There is an entirely new conclusion to the article. The reviewer's comment was well made and has been addressed.