

Elusive quality: how should we evaluate higher education?

BRIEFING PAPER

November 2020

SMF

Social Market
Foundation

By Aveek Bhattacharya, Chief Economist

Monitoring and evaluation of standards in higher education (HE) is in a state of flux, with the future of the Teaching Excellence and Student Outcomes Framework in doubt amidst growing political concern over supposedly 'low quality' degrees. This briefing paper identifies four things the Government in England should do to develop a more coherent and effective system of quality assessment – if that is what it wants.

FOUR TASKS FOR THE GOVERNMENT IN DESIGNING A BETTER SYSTEM FOR ASSESSING QUALITY IN HIGHER EDUCATION

1. **Determine what the purpose of higher education is** and thus what exactly should be evaluated:
 - Narrow economic outcomes
 - A more holistic range of social and economic benefits
 - Student satisfaction
 - Achieved learning gain
 - Effective teaching and provision of resources necessary to learn
2. **Identify the intended audience(s)** and develop assessments appropriate to their needs and preferences:
 - Shaping decisions over funding and accreditation
 - Informing prospective students deciding where to apply
 - Helping employers evaluate workers' qualifications
 - Incentivising and encouraging improvement from institutions
3. **Understand the limitations of quality assessments:** they often come with substantial statistical uncertainty, variation across students and time lags, and so may not support strong judgments.
4. **Decide how important quality assessment is and invest appropriately:** quality assessment should not be done on the cheap – if the Government is unwilling to spend what it takes to properly evaluate higher education, the alternative is to trust institutions more.

Kindly supported by:



INTRODUCTION

Higher education has high stakes. For the Government, it is an investment of £8 billion a year,¹ in what is hoped and expected to be a motor of economic prosperity. For businesses, it shapes the quality of their future workforce. For students themselves, it represents years of effort and dedication, thousands of pounds of debt, and perhaps a route to a better set of life chances. In fundamental ways, it shapes the culture and society we live in. It is understandable, therefore, that so many people have such a strong interest in knowing whether HE is working, and why so many are so concerned and alarmed by suggestions that it is not.

There is growing disquiet, not least in government, about the supposed problem of ‘low quality’ degrees. In May 2019, then Education Secretary Damian Hinds urged higher education providers to “be brave and ask themselves if they’re running courses that really help students gain the skills they need”.² In July 2020, Universities Minister Michelle Donelan suggested that young people had been “taken advantage of”, “misled by the expansion of popular sounding courses with no real demand from the labour market”.³ Such sentiments were reflected in the Augar Review of tertiary education and funding, which recommended “bearing down on low value HE” in the interests of a “small but significant minority of degree students doing certain courses at certain institutions”.⁴ They also led to a commitment in the 2019 Conservative General Election manifesto to “tackle the problem of grade inflation and low quality courses”.⁵

Such concerns have generally been prompted by analysis and modelling of graduate earnings data, which has found that the private financial returns to higher education can be relatively low for a substantial proportion of students. The Institute for Fiscal Studies has estimated that around a fifth will fail to recoup the cost of their degree in terms of higher earnings over the course of their careers.⁶ The Department for Education release that accompanied Hinds’ May 2019 statement referred to the fact that for over one in ten courses, 75% of graduates were earning less than £25,000, and so were below the threshold for making any student loan repayments.⁷ The think tank Onward has used figures such as these to argue that certain subjects and institutions offer a poor “return on investment”.⁸

This perception of systemic failure appears to be behind an ongoing shake-up of the process of quality assessment in higher education. The future of the Teaching Excellence and Student Outcomes Framework (TEF) – the main government initiative in this area, itself only completed for the first time in 2017 – is uncertain. In November 2018, Dame Shirley Pearce was appointed to lead an independent review of the TEF, which was due to report in summer 2019.⁹ Though this is said to recommend substantial changes,¹⁰ the review has still not been published.

In the meantime, the Department for Education has called for a “root and branch” review of the National Student Survey, which underpins the TEF.¹¹ The Destination of Leavers from HE survey, another key data source, has been discontinued.¹² The TEF thus currently exists in a state of limbo. The 2020 iteration was delayed, and although Gavin Williamson, Hinds’ successor as Education Secretary, has expressed his desire for new ratings to be published in 2021, there is some doubt as to whether this will actually occur.¹³

This briefing considers how the Government should proceed in its assessments of higher education quality. Terminology is somewhat inconsistent in this area – the Conservative manifesto worries about “quality”, the Augar review “value” and it is not entirely apparent whether they mean different things.

We are interested here in ‘quality’ in its broadest sense, encompassing ‘value’ (with its connotations of measurable financial outcomes, typically at the individual level), but recognising this as only one of many possible dimensions on which to assess higher education. As the 2013 SMF report *Robbins Revisited* outlines in more detail, the potential benefits of HE are wide-ranging: as well as the individual financial benefits, there are non-financial individual benefits (such as improving health and reducing propensity to commit crime); financial societal benefits (such as higher economic growth and tax revenues); and non-financial societal benefits (such as greater cohesion and social capital).¹⁴ All of these are within scope for this briefing.

The immediate focus here is on policy in England, where the future of quality assessment is currently under debate. However, it is important to recognise that many elements of the existing quality assessment system span the whole of the UK, helping to facilitate mobility of graduates between different parts of the country. Thus some of the ideas and arguments in this briefing will have bearing on devolved HE as well.

This briefing begins with an overview of the existing structures, outlining the roles and functions of the Office for Students, the TEF, the Quality Assurance Agency, Professional, Statutory and Regulatory Bodies, and institutions’ own quality assessment and enhancement processes. It then identifies four things that the Government should do in the process of reforming or replacing these structures:

1. Determine what it thinks the **purpose** of higher education is and thus what exactly it is trying to evaluate;
2. Identify the intended **audience(s)** and develop assessments appropriate to their needs and preferences;
3. Understand the **limitations** of quality assessment in evaluating higher education institutions and courses;
4. Decide how important it thinks quality assessment is and **invest** appropriately.

HOW IS QUALITY CURRENTLY ASSESSED IN HIGHER EDUCATION?

Two main bodies have among their duties monitoring and assessing the quality of higher education: the Office for Students (which operates only in England) and the Quality Assurance Agency (which covers the whole of the UK). For certain subjects, Professional, Statutory and Regulatory Bodies play an important role in accrediting courses. In addition, institutions have their own quality assessment and enhancement processes, most notably the system of external examination. Moreover, with increasing amounts of data publicly available, a number of independent analysts produce their own guides rankings – most prominently newspaper league tables. Further education colleges – where 137,000 people study for HE courses¹⁵ – are also overseen by Ofsted and the FE commissioner, but this briefing will not cover the activities of those agencies.

The Office for Students and the Teaching Excellence and Student Outcomes Framework

How it works

The Office for Students (OfS) is the government-appointed regulator for English higher education. Its objectives are to protect students’ interests by ensuring that their academic experience is of “high quality” and offers “value for money”.¹⁶ In particular, it

emphasises the importance of ensuring students “are able to progress into employment or further study”, and that “their qualifications hold their value over time” (i.e. that ‘grade inflation’ does not occur). OfS’ primary tool for regulation are its conditions of registration, the minimum requirements it imposes on providers to be recognised as higher education institutions, reflecting these overarching objectives.¹⁷ Several colleges and training providers have been refused registration on these grounds,¹⁸ and dozens more, including universities, have been placed under ‘enhanced monitoring’.¹⁹ As of November 2020, OfS has just launched a consultation on how it defines ‘quality’ and ‘standards’ in its regulatory framework.²⁰ Most notably, its proposals include setting numerical benchmarks for the proportion of students that must pass from first to second year, complete a course and get a graduate job in order for a course to retain its registration.

Another prominent way in which OfS pursues its objectives is by carrying out the Teaching Excellence and Student Outcomes Framework in line with the Department for Education’s specification.ⁱ The TEF attempts to assess “excellence in teaching”, as well as the success of institutions in producing positive graduate outcomes for their students.²¹ Universities and colleges are rated as ‘Gold’ (“consistently outstanding”), ‘Silver’ (“high quality”) or ‘Bronze’ (“meeting standards”).²² These ratings are intended to encourage institutions to improve their student experience, and to help prospective applicants choosing an institution to study at.

An institution’s TEF rating is determined by a panel of academics, students and other experts. Ratings draw on three types of data:

- Student satisfaction: how students rate the teaching, assessment and feedback, and academic support they receive.
- Non-continuation i.e. drop-out rates.
- Employment outcomes: whether students are in work or education six months after graduation, and whether the work they find is ‘highly skilled’.

On these measures, providers are judged both in absolute terms, but also relative to a benchmark of how they would be expected to perform given the background of their students and the subjects they teach. There is no set formula or algorithm for converting an institution’s metrics to an overall rating. Instead, the ratings are determined on the basis of the panel’s holistic judgement.

To this point, the TEF has evaluated providers *as a whole*, providing a single rating for the entire university or college. Recognising that teaching, experiences and outcomes may vary substantially between courses at the same institution, there are plans to introduce a subject-level TEF in its next iteration, should it arrive.

Common criticisms

Since its inception, the TEF has been highly controversial. Critics from all directions have argued that it measures the wrong things. Even Professor Sir Chris Husbands, who chairs the TEF panel, admits that “it is not a direct measure of teaching”,²³ and the change in its name (though not its acronym) in 2017 to include ‘Student Outcomes’ reflects a broader focus beyond what happens in lecture halls and seminar rooms.²⁴ That is unlikely to satisfy those that point out that the ‘Teaching Excellence’ component of the TEF relies on measures that are, at best, questionable proxies.²⁵ High student satisfaction could just as

ⁱ When it was introduced, the TEF was initially carried out by the Higher Education Funding Council for England, but responsibility was transferred to OfS in 2018.

well reflect entertaining but educationally ineffective classes, ‘dumbing down’ or low expectations as much as it indicates better teaching or more learning. Worse still, the National Student Survey (NSS), the source for these student satisfaction ratings, has long been dogged by allegations of ‘gaming’, manipulation or inappropriate influence, with issues running from poster campaigns encouraging students to give positive responses to academic staff completing the survey on behalf of their students.²⁶ That is not to say that such views are universal: many providers report finding the NSS an extremely useful source of student feedback and integrate it into their ongoing improvement efforts.²⁷ At the very least, though, its effectiveness as a measure of teaching is contested.

In announcing its review of the NSS, the Department for Education was sharply critical of the survey, claiming that it “has exerted downwards pressure on standards”.²⁸ DfE’s major objection is that “its results do not correlate well with other, more robust, measures of quality, with some of the worst courses in the country, in terms of drop-out rates and progression to highly skilled employment, receiving high NSS scores”. While insisting that student perspectives do have value, the Government would seem to favour a measure of higher education quality that puts greater weight on student outcomes.

The process by which institutional metrics are used to derive TEF ratings has also been challenged. Most notably, in March 2019 the Royal Statistical Society (RSS) wrote to the UK Statistics Authority to express its concern that the TEF “is in large part a statistical artefact”, and as such “is likely to mislead the public”.²⁹ The RSS has raised a number of technical issues that they believe undermine the robustness of TEF ratings.³⁰ They point out that there is substantial statistical uncertainty around many of the metrics that go to make up the TEF, which makes it likely that some institutions will be misclassified on the basis of spurious differences within the ‘margin of error’. This is a particular risk for the TEF because so many different metrics are compared simultaneously, which increases the chances of such errors. As a result, they objected to the apparent precision of classifying institutions as ‘Gold’, ‘Silver’ or ‘Bronze’ in such a clear cut manner. Worryingly, the RSS notes that these sorts of issues are likely to be even more common with subject-level TEF awards, with so many different classifications, based on even smaller samples of data from courses with relatively few students.

The RSS also criticised the lack of transparency around the process by which awards are made. While the underlying data for the TEF is published, the assignment of ratings to institutions depends on judgements which cannot be replicated or scrutinised by outsiders.

The practical value of the TEF in terms of guiding student choice has been questioned too.³¹ In the first year following its publication, 17% of university applicants knew what the TEF was, around 15% used it in some way to guide their choice, and 10% said it was important or extremely important in deciding where to apply.³² However, the publication of the awards had no statistically significant impact on the proportion of applications to gold rated institutions.³³ It is worth recognising, however, that these figures might be seen as a ‘lower bound’, since awareness and influence of the TEF would be expected to grow over time.

Given all these questions regarding its accuracy and usefulness, it is perhaps unsurprising that critics have argued the TEF is not worth the cost and effort that goes into producing it. Central administration of TEF costs £3 million, and there are further costs to participating institutions and individuals (estimated to be around £4 million in total).³⁴ A subject-level TEF would be considerably more expensive: the cost to institutions has been estimated at £15 million, before the central cost of managing 37 panels instead of

just one.³⁵ Such exercises are extremely time-consuming as well: it is notable that the review of the NSS was announced in a government policy paper seeking ways to reduce the “bureaucratic burden” on academics.³⁶

The Quality Assurance Agency

How it works

The Quality Assurance Agency (QAA) exists to “safeguard standards and improve the quality of UK higher education”.³⁷ In England, the QAA has two distinct entities: the ‘Designated Quality Body’, appointed by the OfS to ensure that all academic degrees meet a certain ‘threshold’ standard of acceptability, and an enhancement-focused body to which providers can choose to subscribe. In other parts of the UK, the QAA is funded primarily by subscriptions from universities and colleges themselves, as well as by higher education funding councils.

In contrast to the OfS, its main focus is on institutional processes and structures rather than outcomes. The QAA’s approach is to ensure that providers do what they promise – to ensure students “get the higher education experiences they are entitled to expect”.³⁸ In practice, this involves developing common UK-wide frameworks and procedures on things like staff skills and qualifications, resources and assessments – for example, by developing a Quality Code.³⁹ In its role as designated quality body, the QAA can also go into providers to undertake quality and standards reviews on behalf of the OfS.

It is important to emphasise that the QAA does not judge the standard of teaching or learning *directly*. As the QAA has described it, “We aim to ensure that institutions have effective processes in place to secure their academic standards, but we do not judge the standards themselves”.⁴⁰ As with the existing TEF, the QAA evaluation occurs at the level of institutions as a whole, rather than courses.

Criticisms

There is some disagreement over how effective the QAA is in its role. Its defenders insist that it is genuinely independent, rigorous in its standards, internationally respected and necessary to continuation a common UK-wide HE system.⁴¹ However, it has been criticised for doing too little to ensure that standards are consistent and comparable between institutions and over time.⁴² It cannot, to use a well-worn example, guarantee that a first class degree from the University of Oxford is equivalent to a first from Oxford Brookes University.⁴³ There has also been growing concern in recent years among some analysts over ‘grade inflation’.⁴⁴ Much, for example, has been made of the fact that the proportion of home students at English universities receiving a first has risen from 16% in 2010/11 to 29% in 2017/18.⁴⁵

Those sympathetic to the current system argue that providers’ internal processes of course approval and review (discussed below) are generally sufficiently robust, not least because they are reviewed by the QAA and aligned to QAA expectations. By contrast, critics of the QAA suggest its role is too limited, focusing too much on checking paperwork, and as such is offering insufficient scrutiny. Geoffrey Alderman of the University of Buckingham has described its approach as seeking “to avoid making any judgment about the content of courses...but to insist on a particular form of bureaucratic packaging; this means that a higher value is put on it being absolutely clear and predictable what a student will be told than is put on waking up their minds and seeing how far they can go if they are stretched”.⁴⁶

A 2009 report of the House of Commons Innovation, Universities, Science and Skills Committee agreed with that judgment, and suggested that ideally the QAA would be abolished, were it not for the cost and disruption this would cause. Instead, it recommended re-establishing it as a Quality and Standards Agency with greater independence and a broader remit to more actively assess the quality of teaching and effectiveness of learning.⁴⁷

Professional, Statutory and Regulatory Bodies

For certain HE courses, there is another source of regulation of curriculum content and academic quality: Professional, Statutory and Regulatory Bodies (PSRBs).⁴⁸ These are organisations that regulate entry into particular professions by accrediting related courses. For example, the General Medical Council establishes standards for medical degrees, and the Architects Registration Board does so for Architecture degrees. These bodies are significant because they offer another source of external scrutiny of curriculum, learning outcomes and sometimes facilities and resources. However, given that the requirements of PSRBs vary substantially and are specific to their particular subjects, this briefing will not discuss their role in detail.

Institutions' own processes

How they work

Universities are autonomous institutions given powers to award their own degrees, and so themselves take on much of the responsibility for managing quality. They have a number of processes and structures to protect standards, such as central compliance units that check course materials and learning outcomes are consistent with those of other institutions, by mapping against the reference points in the QAA's Quality Code and qualification frameworks. It is common for courses to go through programme approval, annual monitoring, and more in-depth periodic reviews. Most providers also carry out module evaluation surveys and gather student views for internal purposes.

Internal quality enhancement is increasingly data-led at many institutions. For example, Nottingham Trent University tracks student engagement by monitoring access to online resources, library usage and assignment submissions. Real-time analytics of this sort allows tutors to respond to and address issues as they arise, and as such is extremely promising, though still fairly experimental.⁴⁹

The most prominent mechanism for institutions' own quality assessment, once described as the "jewel in the crown", is the system of external examination.⁵⁰ For many decades, it has been standard practice for academics from outside universities to offer 'peer review' of each course, typically on an annual basis. These external examiners will usually consider the structure and content of the course, teaching materials and assessments. It is common for them to moderate a selection of coursework or exam scripts. They may also meet with students to receive feedback on their experiences.

External examiners usually provide a written report, commenting on the standards of the programme, effectiveness of teaching and assessment, consistency of marking and quality of students' work, and how they compare to general benchmarks in the discipline. The report is then submitted to the institution, which decides whether and how to take action based on the recommendations made by the external examiner.

Criticisms

Though substantial time, effort and resources go into higher education providers' own internal quality assessment processes, the specifics vary from institution to institution, and in general they lack external scrutiny and transparency. The UK's external examination process, under which institutions do invite outside evaluation, was once seen as world-leading, but for some time there has been dissatisfaction with the way it works and suggestions that it is in need of update.⁵¹ In particular, there have been suggestions that it bears some responsibility for apparent grade inflation and the lack of comparability of degrees between institutions.⁵²

Fundamentally, its critics regard it as too informal.⁵³ Examiners are often appointed on the basis of personal relationships. The requirements and obligations of the role are not standardised. There is also a sense that examiners' reports can lack 'teeth', since there is no requirement on universities to publicise or address their findings. Indeed, it has been suggested that external examiners are subtly discouraged from being too critical.

A number of proposals have been made to improve the system of external examination. The Dearing Report, which reviewed higher education in 1997, called for a requirement on institutions to appoint external examiners from a recognised UK-wide pool of independent and specifically trained academics.⁵⁴ The 2009 parliamentary committee recommended that the system be overseen by a retooled QAA, which would specify which documents external examiners have a right to access and what matters they should address in their reports.⁵⁵ A 2011 review for Universities UK and GuildHE endorsed many of these suggestions, but to minimise bureaucracy placed the onus on institutions themselves to be more transparent in selecting examiners, offer greater clarity on their role and provide more support to train them.⁵⁶ Indeed, over the last four years, the Degree Standards Project, an initiative led by AdvanceHE has developed specialist training and delivered it to 2,700 external examiners, forming a directory of trained examiners for institutions to draw.⁵⁷

To ensure that examiners' reports are taken more seriously, it has been suggested that they should be published without redaction and made available to current and future students. It has also been suggested that external examiners should have the power to trigger review of an institution's accreditation to award degrees if serious enough issues are found.⁵⁸

Newspaper League tables

How they work

Though independent of government, and not endorsed by it, newspaper league tables also seek to evaluate the quality of higher education institutions, and potentially exercise greater influence over student applications than official metrics.⁵⁹ The two leading examples, the *Guardian University Guide*⁶⁰ and the *Times Good University Guide*,⁶¹ overlap substantially with the TEF in the measures they incorporate. Like the TEF, both draw on student satisfaction ratings from the National Student Survey, continuation data and graduate employment figures.

However, in addition, they cover a number of features of courses and institutions that are not included by official government measures. Both league tables consider how well students do in their degrees: courses and institutions with more firsts or 2.1s are higher rated (*The Guardian* adjusts these to account for students' level of attainment of school).

They also consider resourcing: universities with lower student-staff ratios or higher expenditure per student do better. Both rankings also reflect the selectiveness of a university: institutions and courses with higher entry requirements are placed higher. Finally, *The Times* ranking includes a measure of research quality, based on the government's Research Excellence Framework.

Criticisms

Newspaper league tables raise a number of issues. First, they can seem rather 'over-stuffed', measuring a range of attributes with limited relation to one another.⁶² Most obviously, *The Times*' inclusion of research quality in its rankings is questionable, since this may have little bearing on the standard of education on offer to students. More generally, combining inputs and outputs, learning outcomes and employment outcomes into a single measure necessitates somewhat arbitrary weightings.⁶³

Some of the specific metrics used are controversial, too. Rewarding institutions for more firsts and 2.1s creates a perverse incentive to relax marking standards, and thus may just reflect grade inflation.⁶⁴ Ranking on the basis of entry tariffs offers an advantage to more established and prestigious providers and courses, regardless of their actual performance and might penalise institutions for taking a more rounded view of an applicant's potential rather than judging them primarily on their grades.

WHAT SHOULD THE GOVERNMENT DO?

As we have seen, there are a range of institutions and an array of metrics that, in different ways, seek to monitor and evaluate quality in higher education. Yet we have also seen that there is widespread dissatisfaction and a belief that the current regime does not adequately identify and address underperformance.

This briefing paper does not offer a fully worked out alternative system – not least because the current frustration often results from unresolved and often unarticulated disagreements over the purpose of quality assessment. Instead, it describes the broad principles and aims policymakers should use when determining how to address these differences and how to go about designing a better system.

1. Determine what the purpose of higher education is

Few questions in higher education are as vexed as what it is for. Yet if we want to understand whether any given institution or course is working well, we have to be clear about what it is trying to achieve.

Improving economic outcomes?

The most obvious answer – and the one that comes back clearest in government rhetoric – is that the purpose of higher education is to improve economic outcomes, both increasing overall national productivity and providing higher incomes and better jobs for graduates. That lends itself to the sorts of outcome measures currently used in the TEF and favoured by think tanks such as Onward: the earnings premium and the proportion of students going into graduate jobs from a particular course. These are not entirely straightforward to specify. For example, how many years after graduation do we need to follow students to understand the impact of their degrees? How do we define a 'highly skilled' job? Focusing narrowly on an individual's earnings may fail to recognise 'spillover' benefits of their career to the wider economy.⁶⁵ However, measures of private economic

gain are relatively familiar and easy to understand, which helps explain their political salience.

The objections to such an approach are equally obvious. Many people see it as excessively narrow and philistine to reduce the value of a degree to its financial benefits. That includes many students: in a recent survey, 84% of students and recent graduates said that future salary was not the only factor in choosing their degree.⁶⁶ Across several studies, students consistently report that they are motivated by passion for their subject and desire for interesting experiences as well as boosting their future career prospects.⁶⁷

Focusing too narrowly on financial gains has a number of problematic implications. For example, it implies that institutions and courses that produce highly paid bankers are much 'better' than those that produce nurses and teachers on more modest salaries. The logic of that would be an HE quality policy that encourages institutions to produce as many future bankers as possible, and as few nurses as possible. Since jobs in London tend to be better paid, such a measure would reward institutions whose graduates work in the capital, rather than in the provinces. Given the gender pay gap, such a metric could disadvantage courses with more female students.⁶⁸ These perverse incentives, if given too much weight, could feed through to recruitment, with institutions less likely to take on students with less social capital, and thus less potential earning power – with extremely undesirable consequences for equality of opportunity.

A holistic measure of outcomes?

One response, then, would be to consider the broader benefits of higher education to students (and indeed wider society) when judging the value or quality of a course. After all, excessive focus on private financial benefits neglects non-financial and societal benefits of the sort identified in *Robbins Revisited*.⁶⁹ That is the approach favoured by Universities UK, a sector trade body which is in the process of developing its own 'charter', providing guidance for institutions as to which metrics to use in evaluating their own courses.⁷⁰ Though the specifics are still to emerge, it has been indicated that these metrics could go beyond personal economic gains to include the 'social value' of the work that graduates go on to do (for example, whether they work in essential public services, in less economically advantaged regions, areas of skills shortage or start their own business).⁷¹ They may also cover non-economic benefits, such as the impact of different courses on life satisfaction, health and community participation.

This approach – producing a long list of desirable benefits from a degree and measuring each institution and course against it – carries major challenges of its own. First, on what grounds is each element included or excluded from the list? Is it, for example, a legitimate objective of higher education to make people healthier or happier? Are those merely side benefits, or indeed, are they irrelevant to 'quality'? Such judgements are highly controversial, and with the holistic list approach, they are myriad. Even if we can agree on a list, operationalising the items on it may be tricky. The concept of a 'socially useful' job is likely to be contested. Measuring health, happiness or community involvement is rarely straightforward.

Student satisfaction?

Given the problems with politicians, bureaucrats or even experts deciding amongst themselves what the purpose of higher education courses should be, there is an understandable impulse to defer to consumer sovereignty. Since students are the ones studying and (in the first instance) paying for the degree, it is sometimes argued that they

should be the ones to decide whether it is any good or not, according to their own values and objectives. This argument lends itself naturally to the use of student satisfaction data of the sort that is integral to the TEF.

As we have seen, the use of such data is highly controversial. While many within the sector see student feedback as a valuable and important indicator of their success, the notion that students are ‘customers’ that are ‘always right’ is strongly resisted by others.⁷² Sceptics argue that satisfaction may reflect low expectations, having an ‘easy ride’, or being dazzled by flashy and inconsequential gimmicks, and as such be entirely unrelated to actual learning.⁷³ One version of this criticism is the paternalistic thought that students cannot in the moment be relied upon to know what is good for them.

A response would be to suggest that we move beyond focusing exclusively on the views of current students as the National Student Survey does, and instead follow up those that have left education to see how they perceive their decision in retrospect. Just as testimony of a car’s reliability carries more weight from somebody who has driven it for a number of years, there might be utility in getting more insight from students several years after graduation.

However, this may still not defuse the objection that students are unreliable assessors of quality. Some may insist that students will *never* adequately appreciate the costs and benefits of their education, and so we cannot rely on their evaluations. There may be an incentive problem with relying too heavily on student assessments: insofar as people benefit from being seen to have attended a ‘better’ institution, it is in their interests to improve the institution’s reputation and thus to overstate its benefits. Crucially, the student is not the only agent with a stake in their education: given public financing of higher education, the government and wider society also have an interest.

Learning, not outcomes?

In any case, there is a deeper objection to evaluating higher education providers on the basis of any set of outcomes, be it economic, student satisfaction or a more holistic measure. Fundamentally, it involves judging them on something that is beyond their control. Consider graduate earnings and employment. They are influenced by a host of factors that have little to do with the actions of providers: macroeconomic policy, labour market regulation, shifting consumption patterns, the business cycle and international trade flows, to name a few. Moreover, these factors may affect different courses and institutions differently. For example, a downturn in a key local industry could, through an accident of geography, appear as a sudden decline in the performance of institutions whose graduates service that industry.

Some might argue that providers should be more ‘entrepreneurial’, and try to anticipate these trends in order to prepare students for the economy of the future. It is important to note that this would involve a fairly significant shift in universities’ and colleges’ role and competences. For all their strengths, there is little reason to think that higher education providers, as currently constituted, are particularly adept economic forecasters who can accurately predict future trends. In pursuing the entrepreneur analogy, it is also worth observing that the market rewards luck as well as skill. It might be better, therefore, to think of positive economic outcomes as reflecting the ‘fitness’ of a course or institution for its economic environment, rather than mistaking this for a measure of its inherent quality.

In which case, it is natural to ask why we do not evaluate what universities and colleges are *actually* constituted to do: impart skills and knowledge. Whatever else students want from a degree, whatever else governments and wider society expect, the thing everyone agrees on is that students are there to learn. That, in turn, suggests that we should evaluate institutions and courses in terms of what people learn from them.

The most straightforward way to do this would be to judge them on the basis of student attainment: how many students pass, and how many get firsts or 2.1s and so on. The problem with such an approach, as we have seen, is that higher education institutions have such wide discretion over degree awards that they are not trusted as reliable measures of attainment, comparable over time or between institutions. Whether or not grade inflation has already occurred, greater emphasis on degree results as a means of evaluating universities would further incentivise it.

If we are to judge HE providers on the basis of learning, we need a more reliable and comparable measure of it. This is tricky because unlike schools, universities do not have a single national curriculum. This brings certain benefits: institutions can be more flexible and innovative with their course content and offer a greater diversity of options for students to choose between. In some subjects, particularly professional qualifications such as Law or Medicine have more standardised curricula, but even these can vary in the way they are taught and assessed in different institutions. One way to provide greater confidence in the robustness of these varied teaching and assessment arrangements would be to enhance the external examination system and put greater emphasis on calibrating standards between institutions teaching the same subjects.

It has also been suggested that there could be more standardised assessments across institutions with, for example, common exam papers or a centrally pooled and administered institution-blind marking system for each subject.⁷⁴ However, in a large and diverse higher education sector, with over 400 providers in England, developing a system appropriate to all or even most providers would be highly challenging to say the least.

It may be preferable to keep separate the process by which universities assess students from the process by which universities themselves are assessed. That would allow universities to keep their autonomy over curriculum and student assessment without having to worry about the knock-on effects for how we judge institutions. There has been growing interest in recent years in developing general or subject specific measures of 'learning gain' that can be used to measure students' progress over the course of their degree. One prominent example was the OECD's Assessment of Higher Education Learning Outcomes (AHELO) project, intended as an HE equivalent to PISA, its standardised test for comparing schooling across different countries. Though the UK government was initially enthusiastic, it ultimately opted against participating in 2015, reportedly due to resistance from elite institutions - a move believed to have undermined the entire project.⁷⁵ At the time, the government justified the decision on the basis that "progressing our current national programme of work on measuring learning gain provides a more efficient and focused approach for this country".⁷⁶ However, in 2018, a report to the Office for Students indicated that a nationally administered learning gain measure, comparable across providers, would face substantial logistical difficulties in effectively recruiting and sampling students.⁷⁷ The OfS claims to be "taking stock" of the findings, but "no further research into learning gain is planned".⁷⁸

Opportunity for learning (teaching and resources)?

As well as practicality, there is a more conceptual objection to the idea that universities' should be judged on their students' learning: that it removes responsibility from the students themselves. Just as a gym or personal trainer cannot guarantee weight loss without effort and commitment from their clients, so universities may argue that they can offer only the *opportunity* for learning, and it is dependent on the students whether or not they take advantage of this opportunity.

This concept of opportunity for learning suggests that the evaluation of universities should focus on the level and quality of the teaching and resources they provide. To some extent, this can be done through collecting the sorts of statistics included in newspaper league tables: student-staff ratios, expenditure per student etc. However, taken alone, these offer little indication of what happens in lecture theatres and seminar rooms – which is what many people will believe is most important.

Universities' own quality management processes – programme monitoring and reviews, student feedback, external examination – to some extent also aim to evaluate teaching. However, these processes lack transparency: it is not always clear to outsiders what they show about current performance, and what steps are being taken to address issues identified. There could be some argument for wider sharing or publication beyond the institution of the evidence generated in these processes.

Some have therefore argued for an inspection-based approach to assessing teaching, along the lines of Ofsted for schools.⁷⁹ Indeed, Teaching Quality Assessments (TQAs), in place from 1992 to 2001, involved a more hands-on approach, with subject-specific reports compiled by academic assessors on the basis of site visits and interviews as well as reviewing paperwork.⁸⁰ Moreover, FE colleges continue to be inspected by Ofsted in the present day, as do some elements of university provision such as initial teacher training and elements of degree apprenticeships. While proponents of the TQA system insist that it was necessary to produce a clear and accurate picture of the state of teaching in higher education, it was strongly resented as intrusive, expensive and unnecessary by the sector.⁸¹ With most departments rated satisfactory, it was also seen as an unnecessary waste of resources.⁸² On the other hand, it could be argued that much of the work of an inspection-based assessment of teaching already occurs through the external examiner system, just without coordination, transparency or publicity.

A note on proxies

Each of these different conceptions of the purpose of higher education – securing economic outcomes, other personal and social benefits, student satisfaction, ensuring or facilitating learning – imply that measures of quality should aim to measure a different thing. However, it could be that it is not feasible, practical or economical to measure the thing we really care about, requiring us to measure something else as a proxy. For example, reported student satisfaction could be taken as a proxy for effectiveness of learning or teaching. In such cases, it is important to try and validate the proxy as far as possible, and to provide evidence that it is in fact a reliable stand-in for the actual outcome of interest. It is also important to be clear and to remember that the measure is only a proxy, and thus to be appropriately circumspect in using it.

2. Identify the intended audience(s) and develop assessments appropriate to their needs and preferences

There are, broadly speaking, four different types of audience for quality assessments, each with different purposes in mind:

- First, they could be used by government, funding and regulatory agencies. In the first instance, they are necessary to determine whether a particular institution or course meets the minimum threshold standards for accreditation and degree awarding powers. However, it has been suggested that 'lower value' courses could receive less public support or be required to reduce their fees.⁸³
- Second, assessments could be used by prospective students and parents to guide their applications.
- Third, they could be used by employers to understand the value of different qualifications when recruiting.
- Fourth, they could be used by providers themselves to better understand how they are performing and how to improve.

These are not, of course, mutually exclusive: the same assessment can be used by more than one of these groups at a time. However, each group has different needs, which the assessment can be better designed to reflect.

The development of quality assessment measures for higher education should, therefore be based on user research and testing to ensure that they are as useful as possible to the intended audience(s). The process should also draw on the evidence we already have on effective performance measures in analogous contexts. For example, parents choosing schools make greater use of Ofsted reports than league tables and statistical indicators.⁸⁴ That offers *prima facie* evidence that inspection reports could be similarly valuable to prospective students.

Each of the four groups is likely to vary in their relative interest in process, as opposed to outcomes. For employers, some students and, to a certain extent the government, it is more acceptable for the assessment to be a 'black box', which identifies the consequences of studying for a particular degree, but not how those results are achieved. As a result, these audiences are likely to have greater interest in measures of financial and broader personal and social benefits. By contrast, for providers, if quality assessment is to trigger improvements in performance it is vital to have as much information as possible on where they are currently falling short. Such detail is more likely to come from more direct evaluation of learning and teaching. Indeed, if improvement is the goal, it would be advisable to have devoted resources and support to help providers act on the findings of quality assessments. Perversely, direct funding for teaching enhancement in higher education has been eliminated entirely since 2018.⁸⁵

Another possible difference is over the question of whether the quality measure should reflect how selective or prestigious a course is – as newspaper league tables do by including entry tariffs. From the perspective of students, the signalling value of the name on their CV is part of the benefit they get from the degree. From the perspective of employers, it offers useful information for a 'pre-selection' process. However, if the government or provider wants to know how effectively a course *improves* its students, it is better to use 'value add' measures that strip away prior attainment and an institution's 'brand'.

Depending on the purpose for which they are used, detail and precision may be more or less important. Being overwhelmed by too much data is likely to be more of an issue for applicants and employers, who may therefore prefer the relative simplicity of the 'gold/silver/bronze' TEF categories. For some regulatory decisions, a binary categorisation of 'satisfactory' or 'unsatisfactory' could be all that is needed. By contrast, for institutions to enhance their performance, more fine-grained judgements across a range of domains are likely to be more useful.

It is of course, always desirable for quality assessments to be as robust as possible. However, the more consequential the judgement, the more important it is to have utmost confidence in it. Arguably, if the purpose is merely to guide student choices, then limitations and caveats are more acceptable. By contrast, if quality measures are to be used to close down departments or restrict their funding, then the burden of proof should be far higher.

3. Understand the limitations of quality measures

This point is critical to recognise, because measures of quality typically have several limitations, and need to be interpreted carefully. This is especially true of the quantitative measures that seem to be favoured by the Government. Three are particularly worth mentioning.

The first, as discussed above in relation to the Royal Statistical Society criticism of the TEF, is that HE statistics are often based on relatively few datapoints, giving rise to substantial statistical uncertainty. The problem is particularly acute when dealing with courses rather than institutions. As a result, it may be impossible to be precise in our rankings or classifications – it may often be the case that providers and courses are functionally indistinguishable from one another, with any apparent difference attributable to random 'noise'.

The second is that averages may be misleading if they mask wide variation in the benefits from a particular course or provider for particular types of student. Students with different backgrounds, prior knowledge or attainment or learning styles may fare quite differently on a particular course. From the perspective of a student making an application, the relevant question is not how a 'typical' student (if such a thing exists) does, but how people like them do. From the perspective of the policymaker, the relevant question might be the effect on the 'marginal' student, if numbers on the course were to expand or contract. This might be quite different from the effect on the average student.

These two points interact: the best way to achieve statistically robust estimates of the effect of a degree is to group students or courses, or even institutions. But the more students are aggregated, the more the differences between them are elided.

The third issue is that assessments and data will often come with a substantial time lag. This issue is greatest for earnings figures: it often takes 5-10 years for the economic value of a degree to become apparent, and sometimes even longer. Yet outcomes for students who studied a particular course ten years or even longer ago may be a poor guide to the effectiveness of the course today, which may have changed in structure and content and may be delivered by different lecturers and tutors to students entering a quite different labour market. Yet even data from direct teaching assessments may quickly become outdated if they are not frequently repeated – as has been the case for many schools' Ofsted ratings.⁸⁶

As a result, it is important for any quality measure to be treated with appropriate caution, and for the government to ensure that judgements and decisions made on their basis are proportionate to the level of evidence they provide.

4. Decide how important quality assessment is and invest appropriately

Some of these limitations of quality assessment are inescapable: courses should not increase in size to provide more reliable data, and without clairvoyance or time travel, we will never with any great confidence be able to predict the futures of individual prospective students. Others, however, can be mitigated with more resources: a government may be able to get more robust quality assessments if it is willing to pay for them.

Progress on national standardised learning gain measures has ground to a halt, because they were deemed too expensive and cumbersome to implement. More direct teaching assessment, in the form of the TQA, was abandoned on similar grounds. The impression given, sometimes, is that the government is trying to do quality assessment on the cheap.

Inspection of schools is seen as a national priority, and so Ofsted has a budget of £130 million a year.⁸⁷ By comparison, the Quality Assurance Agency spent £14 million last year, and is funded by providers rather than government.⁸⁸ The TEF as we have seen, costs around £3 million to administer (though the National Student Survey may be somewhat more), and even that is resented to an extent.⁸⁹ In terms of government spending, school inspections cost around £15 per student, compared to £2 per student for the TEF.

The question, then, is whether it is worth it? The Government might spend more on quality assessment. It might demand more time from and impose more of an administrative burden on academics, diverting them from teaching, research and knowledge exchange. But it should only do so if it believes the knowledge gained is important enough to justify the cost. It may well be that a decision is taken that these costs cannot be justified, especially in the new fiscal context of the Covid-19 economy. In that case, the Government needs to reconcile itself to trusting providers more. In the absence of a more comprehensive – and thus, expensive – assessment regime, policymakers considering analysis and criticism of HE providers and courses should exercise more epistemic humility and recognise the limits of our current understanding of “quality”.

CONCLUSION

Assessing the quality of higher education is not straightforward. It requires deep conceptual thinking and controversial judgements about the purpose of the enterprise. It involves careful consideration about the intended audience for the assessment and their needs and preferences. It calls for adequate epistemic humility, an appreciation of what is feasibly knowable. Done well, it has the potential to ground informed and useful judgements and perhaps to drive improvement and value for money.

Assessing quality is unlikely to be done well on the cheap. A more effective quality assessment is likely to cost money, time and energy. The question for the Government is whether it is enough of a priority to merit such investment.

ENDNOTES

¹ Britton et al (2019), “2019 annual report in education spending in England”, *Institute for Fiscal Studies*,

<https://www.ifs.org.uk/publications/14369#:~:text=Education%20spending%20is%20the%20second,about%204.2%25%20of%20national%20income.>

² <https://www.gov.uk/government/news/education-secretary-calls-for-an-end-to-low-value-degrees>

³ <https://www.gov.uk/government/speeches/universities-minister-calls-for-true-social-mobility>

⁴ Augar et al (2019), “Independent panel report to the Review of Post-18 Education and Funding”, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/805127/Review_of_post_18_education_and_funding.pdf

⁵ The Conservative and Unionist Party (2019), “Get Brexit Done. Unleash Britain’s Potential. The Conservative and Unionist Party Manifesto 2019”, https://assets-global.website-files.com/5da42e2cae7ebd3f8bde353c/5dda924905da587992a064ba_Conservative%202019%20Manifesto.pdf

⁶ Britton et al (2020), “The impact of undergraduate degrees on lifetime earnings”, *Institute for Fiscal Studies*, <https://www.ifs.org.uk/uploads/R167-The-impact-of-undergraduate-degrees-on-lifetime-earnings.pdf>

⁷ <https://www.gov.uk/government/news/education-secretary-calls-for-an-end-to-low-value-degrees>

⁸ O’Brien et al (2019), “A question of degree”, *Onward*, <https://www.ukonward.com/wp-content/uploads/2019/01/J6493-ONW-A-Question-Of-Degree-190104.pdf>

⁹ <https://www.gov.uk/government/news/higher-education-expert-to-review-tef>

¹⁰ <https://www.timeshighereducation.com/news/pessimistic-prognosis-subject-tef-pearce-review-nears>

¹¹ <https://www.gov.uk/government/publications/reducing-bureaucratic-burdens-higher-education/reducing-bureaucratic-burdens-on-research-innovation-and-higher-education>

¹² <https://wonkhe.com/blogs/what-the-extra-tef/>

¹³ Ibid.

¹⁴ Willetts (2013), “Robbins Revisited: Bigger and Better Higher Education”, *Social Market Foundation*, <https://www.smf.co.uk/wp-content/uploads/2013/10/Publication-Robbins-Revisited-Bigger-and-Better-Higher-Education-David-Willetts.pdf>

¹⁵ Association of Colleges (2020), “College Key Facts 2019/20”, <https://www.aoc.co.uk/about-colleges/research-and-stats/key-further-education-statistics>

¹⁶ <https://www.officeforstudents.org.uk/advice-and-guidance/regulation/the-regulatory-framework-for-higher-education-in-england/>

¹⁷ <https://www.officeforstudents.org.uk/advice-and-guidance/regulation/conditions-of-registration/initial-and-general-ongoing-conditions-of-registration/>

¹⁸ <https://www.officeforstudents.org.uk/advice-and-guidance/the-register/refused-registration-decisions/>

¹⁹ <https://wonkhe.com/blogs/what-could-a-registration-refusal-for-a-college-mean-for-everyone-else/>

²⁰ Office for Students (2020), “Executive summary: consultation on regulating quality and standards”, <https://www.officeforstudents.org.uk/media/9b1c68ba-2274-4b71-979c-1fcdf31ae93e/consultation-on-quality-executive-summary.pdf>

²¹ <https://www.officeforstudents.org.uk/advice-and-guidance/teaching/about-the-tef/>

²² Ibid.

²³ <https://wonkhe.com/blogs/tef-results-the-chairs-post-match-analysis/>

²⁴ Department for Education (2017), “Teaching Excellence and Student Outcomes Framework: lessons learned from Year Two”, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/651157/DfE_TEF_Year_2_Lessons_Learned-report.pdf

²⁵ <https://wonkhe.com/blogs/comment-tef-shows-performance-management-is-here-to-stay-but-how-it-works-can-change/>

²⁶ <https://www.timeshighereducation.com/news/national-student-survey-nss-manipulation-claims-raise-questions-about-data-reliability>; <https://www.theguardian.com/higher-education-network/2015/aug/13/the-national-student-survey-should-be-abolished-before-it-does-any-more-harm>

²⁷ <https://guildhe.ac.uk/the-national-student-survey-a-tool-for-enhancing-provision-or-driving-down-standards/>

²⁸ Department for Business, Energy & Industrial Strategy & Department for Education (2020), “Reducing bureaucratic burden in research, innovation and higher education”,

<https://www.gov.uk/government/publications/reducing-bureaucratic-burdens-higher-education/reducing-bureaucratic-burdens-on-research-innovation-and-higher-education>

²⁹ <https://osr.statisticsauthority.gov.uk/wp-content/uploads/2019/03/Letter-from-RSS-to-Ed-Humpherson-TEF.pdf>

³⁰ Royal Statistical Society (2019), “Submission to the Independent Review of the TEF”,

https://osr.statisticsauthority.gov.uk/wp-content/uploads/2019/03/RSS_Evidence_to_the_TEF_consultation_Feb2019.pdf

³¹ <https://wonkhe.com/blogs/so-long-and-thanks-for-all-the-tef/>

³² UCAS (2018), “The Teaching Excellence and Student Outcomes Framework (TEF) and demand for full-time undergraduate higher education”, <https://www.ucas.com/file/173266/download?token=OVbDbdKZ>; Vivian et al (2018), “TEF and informing student choice”, *Department for Education*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/717771/TEF_and_Informing_Student_Choice.pdf

³³ UCAS (2018), op. cit.

³⁴ Universities UK (2017), “Review of the Teaching Excellence Framework Year 2: Process, Results and Next Steps”, <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2017/review-of-the-teaching-excellence-framework-year-2.pdf>; <https://wonkhe.com/blogs/so-long-and-thanks-for-all-the-tef/>

³⁵ <https://wonkhe.com/blogs/so-long-and-thanks-for-all-the-tef/>

³⁶ Department for Business, Energy & Industrial Strategy & Department for Education (2020), op. cit.

³⁷ <https://www.qaa.ac.uk/>

³⁸ <https://www.qaa.ac.uk/about-us/what-we-do/our-work>

³⁹ <https://www.qaa.ac.uk/quality-code>

⁴⁰ House of Commons, Universities, Science and Skills Committee (2009), “Eleventh Report: Students and Universities”, <https://publications.parliament.uk/pa/cm200809/cmselect/cmdius/170/17008.htm>

⁴¹ <https://www.timeshighereducation.com/blog/defence-qaa-private-sector-view>; <https://www.theguardian.com/higher-education-network/blog/2014/oct/20/qaa-uk-university-review-critics>

⁴² Richmond (2018), “A degree of uncertainty”, *Reform*,

<https://reform.uk/research/degree-uncertainty-investigation-grade-inflation-universities>

⁴³ House of Commons, Universities, Science and Skills Committee (2009), op. cit.

⁴⁴ Richmond (2018), op. cit.

⁴⁵ Office for Students (2020), “Analysis of degree classifications over time: Changes in graduate attainment from 2010-11 to 2017-18”,

https://www.officeforstudents.org.uk/media/1e9dacce-8812-4ae6-9216-164e0d817a1d/ofs2019_28_updatefeb2020.pdf

⁴⁶ House of Commons, Universities, Science and Skills Committee (2009), op. cit.

⁴⁷ Ibid.

⁴⁸ https://www.hesa.ac.uk/collection/c12061/psrb_faq

⁴⁹ Jisc (2016), “Learning analytics in higher education”, <https://analytics.jiscinvolve.org/wp/files/2016/04/CASE-STUDY-I-Nottingham-Trent-University.pdf>

⁵⁰ House of Commons, Universities, Science and Skills Committee (2009), op. cit.

⁵¹ Watson (2006), “Who killed what in the quality wars?”, QAA,

<https://wonkhe.com/wp-content/wonkhe-uploads/2019/10/Quality-Matters-December-2006.pdf>

⁵² Richmond (2018), op. cit.

⁵³ Universities UK & GuildHE (2011), “Review of external examining arrangements in universities and colleges in the UK: Final report and recommendations”,

<https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2011/review-of-external-examining-arrangements.pdf>;

QAA & HEA, “External examiners’ understanding and use of academic standards”,

<https://www.heacademy.ac.uk/system/files/downloads/external-examiners-report.pdf>;

<https://www.timeshighereducation.com/features/what-low-quality-course>

⁵⁴ The National Committee of Inquiry into Higher Education (1997), “Higher Education in the learning society”,

<http://www.educationengland.org.uk/documents/dearing1997/dearing1997.html>

⁵⁵ House of Commons, Universities, Science and Skills Committee (2009), op. cit.

⁵⁶ Universities UK & GuildHE (2011), op. cit.

⁵⁷ AdvanceHE (2020), “Degree Standards Project Briefing number six – October 2020”,

<https://www.advance-he.ac.uk/sites/default/files/2020-10/Degree%20Standards%20Project%20Briefing%20No%206%20-%20final.pdf>

⁵⁸ House of Commons, Universities, Science and Skills Committee (2009), op. cit.

⁵⁹ Gibbons et al (2015), “Student satisfaction, league tables and university applications: evidence from Britain”, *Economics of Education Review*,

http://eprints.lse.ac.uk/62875/1/_lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_Gibbons,%20S_Student%20satisfaction_Gibbons_Student%20satisfaction_2015.pdf

⁶⁰ <https://www.theguardian.com/education/2019/jun/07/methodology-behind-the-guardian-university-guide-2020>

⁶¹ <https://www.thetimes.co.uk/article/good-university-guide-2021-methodology-how-we-compiled-the-rankings-j563p2pn0>

⁶² Turner (2005), “Benchmarking in universities: league tables revisited”, *Oxford Review of Education*,

<https://www.tandfonline.com/doi/abs/10.1080/03054980500221975?scroll=top&needAccess=true&journalCode=core20>

⁶³ Ibid.

⁶⁴ <https://www.theguardian.com/education/2007/apr/24/highereducation.universityguide>

⁶⁵ Conlon et al (2018), “Understanding the limitations of graduate outcome measures in higher education”, *London Economics*,

<https://www.guildhe.ac.uk/wp-content/uploads/2018/09/Understanding-the-limitations-of-graduate-outcome-metrics-in-higher-education-18-09-2018-V2.3.pdf>

⁶⁶ <https://www.universitiesuk.ac.uk/news/PublishingImages/Pages/Approach-to-judging-the-%E2%80%98value%E2%80%99-of-a-degree-needs-overhaul/Professor%20Buckingham%20OpEd%20on%20value%20of%20higher%20education.pdf>

⁶⁷ Bartram (2016), “‘Career and Money Aside, What’s the Point of University?’ A Comparison of Students’ Non-economic Entry Motives in Three European Countries”, *Higher Education Quarterly*, <https://onlinelibrary.wiley.com/doi/abs/10.1111/hequ.12098>;

<https://www.timeshighereducation.com/student/news/why-do-students-go-university-and-how-do-they-choose-which-one>

⁶⁸ Cornell et al (2020), “Mind the (Graduate Gender Pay) Gap”, Higher Education policy Institute, https://www.hepi.ac.uk/wp-content/uploads/2020/11/Mind-the-Graduate-Gender-Pay-Gap_HEPI-Report-135_FINAL.pdf

⁶⁹ Willetts (2013), op. cit.

⁷⁰ Universities UK (2020), “Terms of reference for UUK development of charter for enhancing portfolio reviews”, <https://www.universitiesuk.ac.uk/news/Documents/terms-of-reference-charter-for-portfolio-review-november-2020.pdf>

⁷¹ <https://www.universitiesuk.ac.uk/news/PublishingImages/Pages/Approach-to-judging-the-%E2%80%98value%E2%80%99-of-a-degree-needs-overhaul/Professor%20Buckingham%20OpEd%20on%20value%20of%20higher%20education.pdf>

⁷² <https://blogs.lse.ac.uk/highereducation/2020/08/07/when-is-the-customer-right/>

⁷³ <https://wonkhe.com/blogs/can-we-live-without-the-nss/>;
<https://wonkhe.com/blogs/comment-tef-shows-performance-management-is-here-to-stay-but-how-it-works-can-change/>; <https://www.timeshighereducation.com/blog/nss-and-teaching-excellence-wrong-measure-wrongly-analysed>

⁷⁴ <https://www.timeshighereducation.com/features/what-low-quality-course>

⁷⁵ <https://www.timeshighereducation.com/news/oecd-s-ahelo-will-fail-launch-says-education-director>

⁷⁶ Ibid.

⁷⁷ Jones-Devitt et al (2019), “Evaluation of the National Mixed Methods Learning Gain Project (NMMLGP) and Student Perceptions of Learning Gain, *Sheffield Hallam University*,

<https://www.officeforstudents.org.uk/media/6a83a1c6-3933-4d9b-8b95-43f7bfe142e4/final-evaluation-of-nmmlgp.pdf>

⁷⁸ <https://www.officeforstudents.org.uk/advice-and-guidance/teaching/learning-gain/national-mixed-methodology-learning-gain-project/>

⁷⁹ <https://www.timeshighereducation.com/news/ofsted-universities-seen-last-resort-quality-drive>

⁸⁰ <https://dera.ioe.ac.uk/26289/1/QualityStandardsGuide.pdf>

⁸¹ <https://wonkhe.com/wp-content/wonkhe-uploads/2019/10/Quality-Matters-December-2006.pdf>

⁸² <https://www.theguardian.com/higher-education-network/2018/jun/01/awarding-university-subjects-gold-medals-is-deeply-flawed>;
<https://dera.ioe.ac.uk/26289/1/QualityStandardsGuide.pdf>

⁸³ O'Brien et al (2019), *op. cit.*

⁸⁴ Montacute & Cullinane (2018), "Parent Power 2018", *Sutton Trust*,
<https://www.suttontrust.com/wp-content/uploads/2019/12/Parent-Power-2018.pdf>

⁸⁵ <https://wonkhe.com/blogs/after-the-goldrush-30-years-of-teaching-quality-enhancement/>

⁸⁶ <https://www.theguardian.com/education/2018/may/24/ofsted-not-inspected-hundreds-schools-decade-report>

⁸⁷ Ofsted (2019), Annual Report and Accounts 2018-19,
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/817247/CCS0519265192_Ofsted_Annual_Report_and_Accounts_2018-19-WEB.pdf

⁸⁸ QAA (2019), "Annual Report and Financial Statements 2018-19",
https://www.qaa.ac.uk/docs/qaa/about-us/qaa-annual-report-and-accounts-2018-19.pdf?sfvrsn=36b5cd81_4

⁸⁹ <https://wonkhe.com/blogs/so-long-and-thanks-for-all-the-tef/>;
<https://wonkhe.com/blogs/can-we-live-without-the-nss/>