The English Teaching Excellence (and Student Outcomes) Framework: Intelligent accountability in higher education?

Rosemary Deem¹ · Jo-Anne Baird²

Published online: 3 November 2019
© The Author(s) 2019

Abstract

This paper explores what underlies the recent introduction of a Higher Education Teaching Excellence Framework (TEF) in England. Related changes to the higher education landscape are discussed: the 2017 Higher Education Act and creation of a new HE regulator, the Office for Students. How TEF works and some of the consequences of TEF are outlined. As well as discussing what constitutes teaching excellence and what TEF itself is attempting to signal (which includes graduate destinations), we also analyse the underpinning ideologies and logics of choosing metrics to assess teaching excellence, albeit accompanied by peer panel evaluation of institutional written submissions, in determining Gold, Silver and Bronze TEF outcomes. We introduce the notion of TEF as an index rather than a measure. It is suggested that what underlies the English TEF is not about improving teaching but rather an endeavour to pit universities against each other in a highly marketised competitive system with an oversupply of places, in which student debt levels are rising fast. TEF policy is considered with respect to features that would be required of an intelligent accountability system for higher education teaching quality.

Keywords Policy · Teaching excellence · Metrics · Marketisation

Introduction

In this paper we explore what underlies the recent introduction of the higher education (HE) Teaching Excellence Framework (TEF) in England. This includes marketisation of the English system, the 2017 Higher Education Act and creation of a new HE regulator, the Office for Students, and other relevant developments such as Brexit. The paper also looks at how the TEF works and some of the

✉ Jo-Anne Baird
jo-anne.baird@education.ox.ac.uk

¹ Principal's Office, Royal Holloway (University of London), Egham, Surrey, UK

² Department of Education and St Anne's College, Oxford University, Oxford, UK

likely (unintended and intended) consequences of the TEF. In autumn 2017 it was renamed the Teaching Excellence and Student Outcomes Framework (TESOF) but for simplicity we refer to it throughout as the TEF. We discuss what might constitute teaching excellence and what the TEF itself is attempting to signal (including graduate destinations). We also analyse the underpinning ideologies, processes and logics of the TEF exercise itself and explain the thinking behind the choice of metrics to assess teaching excellence, albeit ameliorated by peer evaluation of written institutional self-assessments, in determining Gold, Silver and Bronze TEF outcomes. We note that the metrics chosen constitute an index rather than a causal variable, and we indicate the differences between the two and the consequences. We suggest that the TEF as currently constituted does not increase the value attached to teaching in HE in England, fails to reward academics who are talented teachers and does not necessarily provide students with much useful information about choice of university. Rather TEF is an endeavour to pit higher education institutions (HEIs) against each other in a highly marketised and expensive system with an oversupply of places and rising student debt levels. Indeed, non-repayment of a good proportion of loans is a strong possibility if graduates do not enter well-paid private sector jobs for much of their career. The TEF emphasises competition between institutions far more than collaboration; indeed, if anything, TEF discourages collaboration since each institution is an island. This article contributes to the international literature on tertiary level educational change related to teaching by presenting an English higher education governance and audit change case study and repositioning the use of governance metrics as indices. Assessing the quality of teaching in universities is becoming a hot topic in many countries, from Indonesia (Gaus 2015) to the USA (for example via the established Carnegie Foundation awards for Professors of the Year showing extraordinary dedication to undergraduate teaching) and Australia. Australia is developing measures not dissimilar to TEF but working through the public funding system (Ross 2019). In Europe, there is also a great interest in assessing the quality of teaching, both from the European Commission (2013) and from the European Association for Quality Assurance in Higher Education (ENQA).

The TEF is part of what a report on research metrics (Wilsdon 2015) referred to as the ‘metric tide,’ something underpinned by the development of new managerialism in HE (Deem 2017), and by trying to achieve common HE quality standards across Europe (Cardoso et al. 2015). As a metric, the TEF is a technology for holding HEIs to account. To whom they should be held to account and for what, are important matters that we return to later in this article. Here, we analyse the TEF in relation to the notion of *intelligent accountability*, introduced by O’Neill (2002). Although we also draw on other concepts and bodies of work, intelligent accountability is the central issue. Accountability systems, it had hitherto been argued, promoted trust in modern society because they held professionals to account publicly for delivering products or services, and they improved transparency. With a devastating philosophical critique, O’Neill (2002) showed that accountability systems in fact undermine trust because of the external imposition of targets rather than authentic interactions and dialogue. Through incentive systems, professionals’ attention is faced towards the targets and they are actively dis-incentivised from thinking for themselves about what they might see as genuine improvements. Further,

professionals are kept busy with these externally set targets, so resources for professional engagement and change are depleted by the accountability system.

The expansion of accountability systems accompanies a vision of society in which we have to monitor one another and gauge one another's performances in relation to usually centrally-set targets. O'Neill (2002) pointed out that this is self-evidently not a foundation for a society built upon trust and one which reinforces trust. Fukuyama (1995) considered that accountability systems were put in place *because* trust had dissipated in some modern societies; that they were a reaction to reduced trust. Sahlberg (2010) argued from the Finnish counter-example that accountability mechanisms were not a necessary feature of high-performing education systems, but Finland is a high-trust society and policy borrowing from this context may not work well for other, lower-trust contexts.

O'Neill's (2002) countermove was to propose intelligent accountability systems, which would be based upon self-governance, involve independent judgments as well as internal evaluation, and would not only involve standardised metrics. In later work, O'Neill (2013) proposed that we need to work with genuine units of account, to focus upon quality rather than relative standing, to reflect educational rather than accountability objectives, and emphasised that informed, independent judgments are needed to make intelligent judgments. Further, she noted that gaming of the system will occur if credit is given to less valuable behaviours within the accountability system. Sahlberg (2010) added to this by arguing that collaboration between schools and social networks are important priorities for school systems. Thus, an intelligent accountability system should reinforce these positive educational forces.

We also draw upon a number of other bodies of work, including the use of metrics in HE (Wilsdon 2015) and high- and low-fidelity approaches to assessing teaching excellence (Land and Gordon 2015). We explore different types of HEI with varied stances on matters such as reputation, rankings and quality of teaching and research (Paradeise and Thoenig 2015), and consider whether this relates to the patterns of TEF 2017 and 2018 awards.

Adding to recent contributions to the analysis of the TEF (Barkas et al. 2017; Perkins 2018), this paper emphasises connections between the TEF and other recent changes to the HE system. Finally, we conclude with a discussion about the characteristics of intelligent accountability for improving teaching excellence, but we recognise that this is not the only, or even the main, policy driver for the TEF. This article is important internationally because many countries are considering how to approach the defining and measuring of teaching excellence and there is a long-standing debate about this (Land and Gordon 2015; Elton 1998). England serves as a case for others to study; we observe that HE polices in England have often been borrowed abroad.

Background

The TEF is part of a basket of market HE measures introduced by a Conservative government after the 2015 General Election. Initial, publicly declared TEF purposes were: rebalancing the over-emphasis on research brought about by the Research

Excellence Framework (REF), raising the quality of HE teaching, and providing a basis upon which applicants could choose between HE providers. The TEF was meant to be low cost but a number of factors mean the costs are rising and may soon rival their research cousin, the REF. These factors include the 2017–2018 and 2018–2019 piloting of subject-based rather than institution-wide evaluations; the time taken and expense incurred in making sense of baseline institutional metrics and benchmarks and prepare institutional self-assessment documents; and the running of peer panels at national level to assess TEF submissions. The TEF was originally intended to permit fee increases above the headline fee level, though this was suspended pending the findings of a broader investigation into the costs of HE for undergraduate students in England launched in February 2018 (the Augar Review). So at the time of writing (May 2019) HEIs receive no extra funding as a result of a TEF award.

Research on teaching excellence emphasises the link to student learning, the significance of teams, and different elements of teaching: curriculum design, course content, assessment, pedagogy and leadership (Elton 1998; Gibbs et al. 2009; Ashwin 2015). Some data feeding into the TEF (e.g. graduate employment outcomes and salaries) are not related to teaching quality but rather to social class, cultural capital and degree subject (Behle et al. 2015). Metrics such as student satisfaction with teaching are indicators are not a direct attempt to measure teaching quality (Spooren et al. 2013). If the wrong indicators or metrics are selected, they can drive perverse behavior (Hanson 2000). It is for this reason that we examine the consequences of the TEF and also analyse concerns related to competition versus collaboration, following the concerns of O'Neill (2002, 2013) and Sahlberg (2013).

Recent changes to the UK and English HE system

As already mentioned, the development of the TEF cannot be fully understood without considering other effects upon the recent course of HE in England. National evaluation of research is long-standing in the UK, beginning in 1986. The REF and its predecessor, the Research Assessment Exercise (RAE), have been the subject of much gaming by institutions, from who is excluded to which units of assessment staff are entered in, and the appointment of research stars very close to the census date for staff eligibility (Lucas 2006; Deem 2016). A recent addition to the REF is a considerable emphasis on the non-academic impact of research (Watermeyer 2012). In addition, considerable efforts to reform the REF and stop gaming over who and what is included have been made (Department for Business 2016). The current and previous Conservative governments in England have claimed that research has become too dominant in HEIs and also believe research is being subsidised by teaching, so the TEF is positioned as an antidote to that. A Knowledge Exchange Framework for knowledge transfer work is also underway. Metrics are agenda-setting tools and with all of these possible metric futures, UK academics will have less time on their hands to set their own agendas, research or otherwise.

At the same time the whole system has gradually moved from a quasi-market to a fully marketised system, where competition between HEIs is encouraged well

beyond the number of student places likely to be needed and where even those entering HE to train for crucial public service work, such as nursing, have had their bursaries replaced by loans repayable on graduation subject to an income threshold. The fee levels and rising interest rates for repayment have consequences other than student fees. The UK Competition and Markets Authority has developed a particular interest in HEIs, fining some institutions for malpractice and providing detailed guidance on how programmes of study are to be advertised and what conditions surround changes made after applicants have made their choices. Teaching collaboration within England (as contrasted with international collaboration) does exist in some public HEIs, but is not fostered by exercises such as the TEF. Indeed increasingly in England, HE collaboration seems only to be encouraged in relation to research and doctoral education (which for home/European Union [EU] students tends to break even or make a loss and so is not part of the same marketised, profit-based system as first degree programmes).

In addition, HE in England has attracted a great deal of media interest over the past couple of years. One of the most contentious issues has been that of Vice Chancellors' pay, which, since the inception of high home student fees, has become a particular *cause célèbre*, with extensive media reports and even the occasional resignation of a Vice Chancellor over very high salary levels, often well beyond those of hospital or local government Chief Executives who hold comparable roles. Those high salaries are typically negotiated by remuneration committees of governors, which also have the Vice Chancellor as a member. Governing bodies in England have at the same time gone heavily down the 'boardism' route (Veiga et al. 2015), with considerable efforts made to recruit business people and professionals such as lawyers and accountants to governing bodies and with governing bodies increasingly encroaching on decisions about academic matters previously the prerogative of Senates or Academic Boards (Deem and Magalhaes 2017).

The 2017 Higher Education Act was the first piece of English legislation on HE regulation for 25 years and brought into being both the Office for Students to regulate and 'police' English HEIs and a new umbrella body for the UK Research Councils, UK Research and Innovation (UKRI). At the same time, since the June 2016 EU referendum vote, UK HEIs have been dealing with the uncertainty surrounding a no-deal Brexit or a poor deal Brexit (Corbett 2016) on matters such as recruitment and retention of EU students and staff (the latter are a significant component of HEI academic staff), loss of EU research funding (the UK has historically had a significant share of such funding) and the Erasmus + student exchanges (Hunter and De Wit 2016; Mayhew 2017). There are likely to be greater difficulties in obtaining visas for international staff, students and attendees at UK academic conferences.

Thus we can see that the TEF is a policy development occurring alongside other significant changes to English and UK HE more widely. The TEF sits alongside the post-Brexit referendum context for the UK, as well as continued government Home Office insistence that international students form part of net migration figures, giving the interpretation of TEF results in the overseas student market a very different emphasis now. However, in September 2019 it was announced that international students would be able to stay and work in the UK for up to 2 years. Nevertheless for EU students, early in 2019 the current Conservative government made clear that

post-Brexit EU student fees will be significantly raised, possibly as early as autumn 2020 (Adams 2019), presumably to the same level as other international students. Only Scotland of the four UK countries has decided not to go down that route but to retain free EU tuition.

Theory, methodology and sources

We draw on relevant academic literature, analysis of TEF policy documents and reports and media coverage. Theoretically we not only use the concept of intelligent accountability (O'Neill 2002), as already outlined, but also utilise work on how teaching-related metrics operate and what they measure (Hanson 2000; Spooren et al. 2013); the underpinning political and ideological process by which metrics are selected; and debates about how teaching excellence can be recognised and rewarded (Elton 1998; Gibbs et al. 2009; Ashwin 2015). Teaching excellence is challenging to assess; it needs, as Elton (1998) noted, to be linked to student learning and is often a team rather than an individual effort. The latter point was also made by Ashwin (2015). Ashwin argued that no-one is excellent all the time and that matters such as curriculum design are critical in excellent teaching. Land and Gordon (2015) argued that teaching excellence initiatives may be either low or high fidelity, with or without significant evidence. Student nominations for teaching excellence are often lacking in evidence beyond liking the people concerned, but awards such as the UK National Teaching Fellowships require considerable evidence and use external referees. Gibbs et al. (2009) focused on how Heads of Department can play a key role in ensuring high-quality teaching by being good role models themselves, listening to staff and students, praising good work, encouraging innovation and being open to change.

We also draw upon research about how different ideal types of HEIs approach academic quality (Paradeise and Thoenig 2015), including teaching. French academics Paradeise and Thoenig undertook a detailed study on academic quality (both research and teaching) in HEIs. They used observations, interviews, documentary analysis, academic CVs and academic outputs in 17 HEIs in six countries—France, Italy, Spain, Switzerland, China and the USA—developing a fourfold categorisation of HEIs in relation to academic quality which ranges from ‘Top of the Pile’ and ‘Venerables’ to ‘Wannabees’ in the middle and ‘Missionaries’. The typology is based on two axes: attention to reputation and attention to excellence. ‘Top of the Pile’ applies to institutions where the importance of both reputation and excellence are of foremost importance. Such HEIs are typically at the top of world research rankings, such as Shanghai Jiao Tong University, and exceptionally good at teaching too. ‘Venerables’ have high reputations and status but do not usually have high marks in world rankings. ‘Venerables’ pay a lot of attention to status and reputation but very little to external endogenous excellence, whose trappings they regard with disdain. They are described by Paradeise and Thoenig as behaving like an aristocracy. ‘Wannabees’ may be newer institutions or former ‘Venerables’ or ‘Missionaries’. Their main aim is to bolster their national and international excellence by paying great attention to excellence and trying to make use of external agencies and

so-called objective quantitative measures in order to raise the level of that excellence. Reputation is of lower concern. ‘Wannabees’ may be absent from or low in international rankings because of factors such as small size, lack of attention to publishing, not enough international staff, focusing on teaching and textbook writing, and so on. Finally, ‘Missionaries’ are the largest group of the four, they are part of the massification process of HE and closer to the polytechnic, *Hochschule* and vocational training tradition than the traditional academic university. Many of these were upgraded to universities in the late decades of the twentieth century and some are located in regions previously without any HE provision. ‘Missionaries’ do not have a consistent attitude to quality. Some have started to move away from teaching intensity whilst others stay there quite happily. For many ‘Missionaries’ international league tables are irrelevant, but others seek excellence in niche fields. When we move onto exploring what has happened in the Teaching Excellence Framework we will return to see how well this typology works with the data on those achieving Gold in TEF 2 (2017) and TEF 3 (2018).

Finally, we make use of the concept of unintended consequences of social action, to speculate on the effects of TEF. Drawing on the work of Krücken (2014), there is a development of Merton’s (1936) work on the unintended consequences of what he called ‘purposive social action’. Krücken observed that Merton’s (1936) work refers to social actors rather than organisations. Hence Krücken reshapes Merton’s approach to take into account organisational actions as the ‘idea of a discursive field in which remarkable change processes take place’ (Krücken 2014), thus allowing the concept of unintended consequences to be applied to organisational contexts. Merton put forward five causes of unintended consequences: error, ignorance, immediate interest, basic values and self-defeating prophecy. At least three of these seem relevant to the TEF: ignorance of what teaching excellence means outside the world of elite private schools such as Eton and tutorial-based teaching at Oxbridge (Oxford and Cambridge universities) with which the Conservative cabinet is so familiar; the immediate interest of the English political establishment in creating, for campaigning purposes, a new exercise about teaching excellence at a time when the introduction of high fees is causing students and their families to question whether taking a degree is still worthwhile (Brown and Lauder 2012) and values about marketised HE systems driving the whole approach to HE as a series of public and for-profit HEIs competing with one another.

What is the Teaching Excellence Framework and what is it for?

The formal launch of the Office for Students in 2018 was enabled by the 2017 Higher Education Act. The Office for Students is a semi-replacement body for the Higher Education Funding Council for England (HEFCE) and is responsible for the TEF. All HEIs in England who want to access public funding for teaching (including student loans) or research (via UK Research and Innovation or UKRI) must be registered with Office for Students. Unlike HEFCE, which was both a regulatory body for HE in England and a buffer between the government and HEIs, the Office for Students is purely a regulatory body and its approach is already evidently different

from that of HEFCE. Student value for money, fair access, positive outcomes for all students, preservation of free speech and treating all institutions the same are some of its main concerns. The Office for Students's main means of regulating HEIs is through fining them. For example, Hertfordshire University was fined recently for overcharging for franchised courses (McKie 2018). Some HEIs have been questioned by the Office for Students about how many unconditional offers they give to students who have yet to sit their final exams, and potentially a number of HEIs could be fined for giving too many first-class degrees (Savage 2019), although it is not clear if the last-named action by the Office for Students is legal. The Office for Students seeks to differentiate itself from HEFCE, its predecessor body, and does not see itself, for instance, as having a role in ensuring public HEIs do not go bankrupt; something which reportedly is a serious prospect for a number of current, leading institutions.¹ Instead, all Office for Students-registered institutions must have a plan in place to deal with institutional failure.

The TEF was first mentioned in the 2015 General Election Conservative Manifesto. Two consultation papers on the TEF appeared in 2015 and 2016 (Department for Business Innovation and Skills 2015, 2016). There was remarkably little resistance from Vice Chancellors, but more from academics. It was suggested by the Department for Business, Innovation and Skills that that the exercise would not be bureaucratic, would not be a burden on HEIs and would help both HE students and their teachers. The first sponsor of the TEF, Jo Johnson (then Higher Education Minister for England 2016–2018 and 2019), saw its main purpose as raising the profile of teaching in relation to research and to inform students and HE applicants about the quality of education and graduate employment outcomes. TEF 1, in 2016, was essentially a desk-based exercise in which all HEs in the UK who had completed an institutional audit with a positive outcome within a given period were allowed to proceed with a small fee increase. TEF 2, in 2017, involved National Union of Students undergraduate survey data; data on the employment of graduates from the sixth months after graduation survey Destinations of Leavers from Higher Education (DLHE), a great deal of benchmarked data and metrics on the achievements of different categories of students (determined by type of institution and student body composition) from those with disabilities to gender and black and minority ethnic students; but also a detailed institutional self-assessment statement.

Though it is primarily an English initiative, institutions from the other three UK countries were allowed to enter. Apart from in Wales, this offer was not taken up very much. No institution, however, was compelled to enter TEF 2 or TEF 3. This will also be true of TEF 4 in 2019, which is still assessing excellence at institutional level. When subject-level TEF is introduced, currently planned for 2020, all public HEIs in England with over 500 students will be *required* to enter. Both reputation and the promised ability to raise home/EU fees beyond the level of inflation were incentives for TEF 2 and TEF 3 at a time when there was a demographic fall in the number of 18–20 year-olds in the UK entering HE. TEF 2 made National Student

¹ <https://www.theguardian.com/education/2018/dec/11/struggling-uk-universities-warn-staff-of-possible-job-cuts>.

Survey (NSS) data from final year undergraduates a significant part of the survey metrics but a number of HEI Student Unions joined a boycott of the survey in spring 2017 as part of a protest about the NSS being used for TEF purposes. Hence in TEF 3 the part played by NSS data was downgraded despite the fact that these data do relate to teaching, whereas graduate employment and earnings data do not. Low response rates to the NSS also make its use problematical.

On 7th September 2017, speaking to Universities UK (UUK)'s annual conference, after TEF 2 had been completed, Jo Johnson said:

The TEF is already transforming learning and teaching across the HE sector, with, for example, Imperial's vice-provost for education describing it as a 'godsend' for teaching in our system. I'm pleased UUK's comprehensive survey of providers found that: 73% believe that TEF will raise the profile of teaching and learning in universities [and] 81% have undertaken additional investment in teaching, with almost half saying the TEF had influenced their decision to do so. And there is general confidence that the overall process was fair (Johnson 2017)

It may have seemed premature to declare TEF successful so soon after its full implementation but that is politics. Politicians usually cannot or do not want to wait for the long-term view because electoral cycles are short.

By January 2018, following a snap election in June 2017 intended to produce a larger majority for the Conservatives, so as to strengthen the hand of Prime Minister Theresa May in the Brexit negotiations, Jo Johnson's star was waning. The June election left the Conservatives without an overall majority. In the winter of 2017–2018, a controversial appointment was made by Johnson of a very right-wing and controversial figure—Toby Young—to the Board of the Office for Students. Young had made lewd comments about women, including members of parliament, on social media and also spoke disrespectfully about people with disabilities and working class students. Following public outcry, Young resigned and Johnson was replaced in January 2018 as Higher Education Minister by Sam Gyimah, who himself then subsequently resigned on 30th November 2018 over the Prime Minister's handling of the Brexit strategy. In December 2018, another new Higher Education Minister was appointed, Chris Skidmore, who launched a review of the TEF on 18 January 2019 chaired by Shirley Pearce, former Vice-Chancellor of Loughborough University. In July 2019, Jo Johnson was re-appointed by the new Prime Minister, his brother, Boris Johnson but he resigned unexpectedly on September 5th 2019, citing an irresolvable tension between his family and the national interest. Gavin Williamson was appointed to replace Jo Johnson. At the time of writing he had yet to demonstrate what his focus and vision for HE would be.

One of Gyimah's first actions in early 2018 had been to announce that institutional TEF would in time be replaced or stand alongside a subject-based TEF. There was a public consultation on this and the pilots for subject-based TEF were trialed during 2017–2018 and 2018–2019 though neither of the initial models used in the 2017–2018 trial were found to be satisfactory. It is not clear yet exactly how subject-based TEF would operate but it is intended that any institutional TEF awards would only last until subject-based TEF starts. Other changes to institutional TEF

were made for the 2018 TEF round 3, which remained an institutional level activity, but in any case had fewer than 25 applications from HEIs as contrasted with well over 100 for TEF 2. TEF 3 changes included giving less credit to NSS survey results and using a new post-graduation survey of graduate job destinations via not just DLHE (graduate employment 6 months on) but also Longitudinal Education Outcomes (LEO) which focuses on rather longer term graduate destinations and salaries than the DLHE. We await what TEF 4 will bring. Current TEF results will last until 2020.

There were signs that Skidmore's ministerial role in HE and his vision for TEF differed from those of both Johnson and Gyimah. In a speech to the Royal Academy of Dramatic Art in January 2019, a few weeks after his appointment, Skidmore said:

Although I appreciate the TEF has raised questions, no university should shy away from it. The independent review of the TEF, which launched earlier this month and is chaired by Dame Shirley Pearce, provides an important opportunity to take stock of the TEF from a constructively critical perspective. As part of the review, I am pleased to note that Dame Shirley has commissioned the Office for National Statistics to carry out an analysis of the statistical information used in TEF assessments and its suitability for generating TEF ratings ... As much as I see the value of more data, I am also aware of concerns it has given rise to about the value for money of certain courses, disciplines and institutions. On this, I believe we need to take a step back and ask what exactly value for money means in the context of higher education. Successful outcomes for students and graduates are about much more than salary: if we are to define value purely in economic terms, based on salary levels or tax contributions, then we risk overlooking the vital contribution of degrees of social value, such as Nursing or Social Care, not to mention overlooking the Arts, Humanities and Social Sciences – the very disciplines that make our lives worth living.

At the time of writing it was too soon to tell what the September 2019 appointee to Education Secretary and Chris Skidmore, reappointed Universities Minister, would do but Williamson appeared to be pushing forward subject-based TEF. The announcement about international students being able to work in the UK for 2 years came too soon for either to have had any significant input. Subject based TEF remained uncertain in autumn 2019.

Measurement has been an incredibly important tool for the sciences and the social sciences. Metrics, or measurements of key data, have helped us to construct models of how things operate and to evaluate interventions to improve them. They are, of course, not just a tool for science. The current tide of metrics in HE does not stem from academics' desire to better understand how research, teaching and impact are operating. Instead, the push has come from a managerialist agenda to hold public bodies more accountable, ensure value for money and improve quality. New public management sits within a neoliberal worldview, which has free market economics, competition, de-regulation and macro-quantification as its founding blocks and new managerialism is an ideologically driven version of this (Deem and Brehony 2005), common in education, health and public services for some years.

From such traditions arises England's Office for Students and its associated metrics for measuring, monitoring and ultimately improving the student-customer experience (Department for Business Innovation and Skills 2016). The Office for Students is a full member of the audit society (Power 1997).

The TEF is not a measurement: it is an index

Really measuring the quality of teaching in HE is a complex task, especially as agreed definitions of teaching quality are not available. Under these circumstances, a tactic can be to turn to 'indices' instead of 'causal variables'. Let us pause for a moment to consider the implications of this fundamental move. A causal variable measure can be made up of more than one aspect (such as a mathematics school-leaving examination), but the assumption is that the underlying variable of interest (e.g. proficiency in mathematics) causes the outcomes in the measure overall (mathematics grade) and its composing parts (examination questions). Such causal variables are often termed the 'construct' being measured in assessment terms. If we wanted a measure of teaching excellence, we would have to select items and an overall outcome that was caused by teaching quality. Factors other than teaching quality affecting that outcome would be construct-irrelevant nuisance factors that invalidate the measure, and the course of action to improve the measure would be to weed them out. For example, if we knew that there were biases in student evaluations of teachers (MacNell et al. 2014; Subtirelu 2015) then we would not use student evaluations or would adjust for such biases to counteract them. For a measure, we would expect the composing parts to correlate well and would question the validity of any aspects that did not meet this criterion. Clearly, for a causal variable measure, the single, often latent, construct of interest defines the quality criteria. Figure 1 indicates in diagrammatical form what this would look like for the TEF. Teaching excellence, as a causal variable would be producing effects in the individual measures that we have such as the NSS regarding student satisfaction, students' labour outcomes in the form of wages, the amount of academic support given and the upholding of academic standards (as opposed to grade inflation). Teaching excellence is postulated to exist and to be a cause of these other manifest variables.

In contrast, an index is a compilation of factors that are of interest in themselves (Tesio 2014). In this case, the causal connection is from the index items to the underlying variable of interest. The items define the variable of interest. A single index is produced which summarises the multiple attributes of interest. An example of such a variable is the consumer price index.² It is composed of a basket of items that might not necessarily correlate, but they would not be rejected from the index on those grounds alone *because* they are of interest in themselves. The World University Rankings produced by the *Times Higher Education Supplement* or the QS World University Rankings are examples of indices. Research volume, research

² <https://www.ons.gov.uk/economy/inflationandpriceindices/articles/ukconsumerpriceinflationbasketofgoodsandservices/2018>.

Fig. 1 TEF as a causal variable

Fig. 2 TEF as an index variable

income, citations, industry income and reputational measures from other HEIs or domestic employers are included in the basket of items considered. Thus, constructors of indices have to be transparent about the content of the basket if we are to have confidence in the quality of the scale. Quality is decided through reference to external factors, in terms of what the index predicts, rather than through internal statistical analysis of the data. Once constructed, if we found that such indices put

the world's most renowned universities at the bottom of the list, the course of action would be to look again at the basket and either add items or remove the offending items so that a sensible rank order was produced. Here we see a quality criterion—whether the results have face validity. Figure 2 shows TEF as an index variable. The measures we have, such as student satisfaction, subsequent wages of students, levels of academic support and grade inflation combine to produce our metric of what we have named teaching excellence. Teaching excellence, in this approach, is something that is defined and constructed by the constituent parts of the index. This is what we call it and though there may be a different definition in use elsewhere, it does not pertain to the rules and regulations that TEF is party to.

The TEF could be seen as an index, then, rather than a direct attempt to measure a causal variable of teaching quality. Customer satisfaction could be seen as an indicator, even if it does not directly measure the quality of teaching. Equally, employment prospects and outcomes *can* be seen as an indicator of teaching quality. We might all accept that other factors *cause* employment prospects and earnings, but remember that an index approach essentially defines the variable of interest. The TEF sets out what HEIs should be achieving; the fact that it is termed 'teaching excellence' is secondary and the nuances of what teaching excellence should look like are moot, because academics cannot agree on how it should be measured in any case. The TEF stance sets out to define a construct of teaching excellence through the measures in the basket.

What is in the TEF basket?

The third round of TEF in 2018 had a number of items (degree grade inflation and longitudinal employment outcomes data) added to the basket (Table 1). As with all neoliberal metric systems, the counter argument to considering the items included in the basket can appear outmoded, inefficient and lacking in a practicable alternative. Who could argue that the items in the basket should *not* be part of a well-functioning HE system? Shouldn't students know when they are getting a good education? Shouldn't prevention of grade inflation be a social good? Aren't academic and pastoral support important? Wouldn't a good education lead to well-paid jobs or further study? An essential argument that is common to neoliberal metric systems is that there is no other alternative, which moves the discussion on to the details of how things are done.

The TEF is a conglomeration of available statistics. The NSS is notorious for its low response rate and difficulties interpreting students' ratings. NSS data are included in the TEF so long as the response rate is 50% within a particular subject field. Although this is not a low response rate for surveys in general, its use in monitoring an institution changes the interpretation of this response rate. Surveys in the spring term of a bachelor's degree are vulnerable to events that take place in that time period (such as building works, strikes, a bad exam or essay experience) rather than the experience over the whole time spent on the programme being evaluated. Further, student perceptions may be quite different from those who are familiar

Table 1 The Teaching Excellence Framework 3

Aspect of quality	Data	Source	'Provider' narrative on all aspects of the TEF
Teaching	1. Student satisfaction on assessment and feedback 2. Grade inflation	1. National Student Survey (NSS) ^a 2. Provider declaration	
Learning environment	1. Academic support 2. Student retention	1. NSS 2. Higher Education and Statistics Academy (HESA) Individualised Learner Record (ILR) data	
Employment or further study	1. Employment 2. Highly skilled employment 3. Further study 4. Sustained employment 5. Above median earnings	1. Destination of School Leavers from Higher Education data from the Higher Education Funding Academy (HEFCE) ^b 2. Longitudinal education outcomes (LEO) from merged government department data ^c	

^a<http://www.hefce.ac.uk/lt/nss/results/>^b<http://www.hefce.ac.uk/lt/dlhel/>^c<https://www.gov.uk/government/statistics/graduate-outcomes-for-all-subjects-by-university>

with a breadth of provision in HE. For example, Oxford and Cambridge University students rate the feedback they receive just as badly as those in other institutions, despite the fact that they receive one-to-one tutorials as their main mode of teaching, when other students are taught in large lecture groups.

‘Grade inflation’ is simply an increase in outcomes from the year 2008–2016. No attempt is made to investigate whether changes in outcomes are warranted by students’ performances. Including this aspect in the index counters one potential way in which HEIs might have gamed the TEF. As student satisfaction is a feature of the TEF and some research shows that higher grades increase student satisfaction (Isely and Singh 2005; Langbein 2008; McPherson 2006; McPherson and Todd Jewell 2007), institutions might have been tempted to game the system by giving students better grades. Indeed, research in the US has suggested that using student satisfaction as a metric has put pressure on academics to give better grades (e.g. Stroebe 2016). Thwarting this gaming strategy by focusing upon increases in outcomes is an interesting tactic that could address unintended consequences. However, it is a curiosity if the object of the TEF is to drive up teaching standards. Improved teaching *should*, after all, improve students’ performances and therefore their degree outcomes. At this stage, the Office for Students are not setting limits on increases in degree outcomes. Narrative explanations can be provided by institutions in relation to grade inflation as well as other data. The direction of travel is set, though, and we can anticipate a formula for grade inflation in future years.

The examinations regulator in the UK, Ofqual (2014), already have a policy (‘comparable outcomes’) to tackle both grade inflation and to thwart perceived competition between exam boards that was feared to dumb down standards (Education Committee 2012). This approach is based upon a body of research (Baird 2007; Baird et al. 2000; Cresswell 1996; Christie and Forrest 1981; Coe 1999, 2007, 2010; Newton 1997a, b, 2003, 2005, 2010) and assumes a standard value added at population level for students taking the same subject qualification (e.g. A-level Physics), even if their prior qualifications were issued by different examining boards in England. Ofqual’s methodology has led to a leveling of the school-leaving examination outcomes (A-level) since its introduction in 2010. However, the proportion of first class degrees has risen from 18% in 2013 to 26% in 2017.³ Although some of the students will be from overseas and therefore will not necessarily have taken the A-level examination, it is unlikely that HEIs have been able to transform their teaching to add such significant value over this four-year period. Examination boards who grade the school-leaving examination are separate from the schools and colleges who teach the pupils. HEIs both teach and grade their students. Different institutions may well consider that they produce differential value-added effects, so a distinctive approach may be required. However, having value-added benchmarks both between and within institutions may be a useful place to start.

Unlike grade inflation, HEIs do not control labour market destinations of their students. HEIs have a role in supplying human capital, but are not responsible for industrial strategy. Further, labour market inequalities cannot be addressed by HE

³ <https://www.gov.uk/government/news/new-measure-to-tackle-grade-inflation-at-university>.

alone. As the Social Mobility Commission Report (2017)⁴ showed, there are certain highly-paid occupations for which social capital (including attending a private school) is a dominant factor (Milburn 2012). Notwithstanding, LEO data on recent graduates (1, 3, and 5 years after graduation) is now being published for institutions and different programmes, purportedly to allow students to work out which courses will lead to higher earnings. Inclusion of metrics related to future earnings clearly needs to take account of sector averages, public and third sector versus the private sector, the gender pay gap and many other things.

Benchmarks are being constructed for many of the TEF metrics, which take into account factors that HESA (formerly known as the Higher Education Statistics Agency; experts in HE data and analysis) recognises may be beyond the control of the institution. These are the subjects of study, entry qualifications, age on entry, ethnicity, sex, disability, level of study, a measure of socioeconomic status and year of undergraduate study. Despite explanations, it is unclear exactly how these benchmarks are compiled.⁵ The benchmarks are not common to all institutions as it depends on the characteristics of each HEI's intake. Even if one HEI's data do not change, the benchmarks could alter in subsequent years, due to changes in other institutions.

The consequences of an index approach to the TEF

If the wrong indicators (or metrics) are selected with which to gauge teaching excellence, they could drive the wrong behaviours in HE organisations. As Hanson argued, the signifier can assume a life of its own (Hanson 2000). Metrics are booming due to the spread of ideas through national and supranational organisations such as the Organisation for Economic Cooperation and Development and the EU. In the 90s, only half of all European countries had external quality assurance arrangements for their HEIs, but since 2003 almost all of them have (Cardoso et al. 2015). Both the Bologna Declaration (1999) and the European Quality Assurance Register for Higher Education have been important agendas in this regard. But the view is that teaching is only measured by economic outcomes. Through those metrics we are coerced and controlled and at times rational actors seek to corrupt the system for a range of reasons. What is missing from the metrics is important as it is unlikely that the TEF alone will lead to improved teaching. But the point of the TEF, of course, is to construct a competitive market for HE along these lines, not to improve teaching per se.

Despite efforts by HESA, which collects UK HE institutional data, the TEF is opaque. It is a bureaucratic, self-referential metric with its own criteria of technicalities and jargon of absolute values, thresholds, flags, benchmark, splits, z-scores,

⁴ https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/662744/State_of_the_Nation_2017_-_Social_Mobility_in_Great_Britain.pdf.

⁵ <https://www.hesa.ac.uk/data-and-analysis/performance-indicators/benchmarks>.

materiality, stars and exclamation marks. The system can be seen in a 90-min YouTube webinar by HESA.⁶

The TEF process

There is a sizeable literature on definitions of teaching excellence (Ashwin 2015; Gibbs et al. 2009; Robinson and Hilli 2016; Wood and Su 2017; Elton 1998) Land and Gordon in the early stages of the TEF debate, considered different approaches to teaching excellence initiatives (Land and Gordon 2015), noting both high- and low-fidelity approaches (with and without verified evidence). Two additional critical commentaries on Land and Gordon's work (Deem 2015; Tsui 2015) were also commissioned by the Higher Education Academy. Since then much, if not all, of the teaching excellence debate in England has taken place outwith formal consultations, which rarely encompass fundamental issues of policy rationale.

There is as yet, little written by TEF panel members reflecting on the TEF process; this is unlike the situation in relation to the Research Assessment Exercise and the newer REF where there are many papers written by former panellists. There is a review of TEF 2 from UUK (Universities UK 2017) and the Higher Education Policy Institute (Beech 2017) and also a report from the Department of Education (the Ministry) which used surveys of key stakeholders, including Vice Chancellors and students involved in provider submissions, as well as panellists and institutions which did not submit (Department for Education 2017). But none of those publications has been through a peer review process and are invariably more positive and upbeat than critiques in journal or conference papers. The UUK study collected views from its members (heads of UK universities) suggesting in a fairly balanced way that:

There appears to be general confidence that overall process was fair... Judgements were the result of an intensive and discursive process of deliberation by the assessment panel.

- The TEF produced independent results that are partially corroborated by other metrics. The results did not correlate with institutional characteristics such as student population or research income, but there was a slight correlation with entry tariff and other rankings.
- Further consideration will need to be given to how the TEF accounts for the diversity of the student body, particularly part-time students.
- Some regional patterns also suggest that the sensitivity of judgements to contextual factors may also need further consideration. ...
- There is widespread belief that the TEF will raise the profile of teaching and learning.

⁶ <https://www.youtube.com/watch?v=bgSETr48Z-A&feature=youtu.be>.

- There is also early evidence that the TEF process has enhanced engagement with institutional metrics, will reshape internal assurance processes and has influenced teaching and learning strategies...
- There are genuine concerns about how the assessment framework defines and measures teaching excellence and the viability of subject level assessment.

The UUK report does comment adversely on the £4 million cost of TEF 2 (for submitting institutions alone). The HEPI report by Beech (2017) contains a foreword by Chris Husbands, Vice Chancellor of Sheffield University and Chair of the first TEF panel which emphasised the good points of the exercise:

the interpretation of core and split metrics, the analyses of contextual material and of the 15-page provider statements were all consistent with the TEF specification, were robust and thorough and were conducted with professionalism and integrity. It was an enormous privilege to chair the TEF, to lead assessors and panellists and to work with the quite exceptional professional support team at the Higher Education Funding Council for England (HEFCE). While there have been other assessments of teaching quality in higher education – such as the subject review process two decades ago – it has never before been attempted at this institutional level, nor at this scale. In short, the running of the TEF was a massive technical challenge. I have worked in higher education for long enough, and in enough different roles, to know that there are always critics, yet the running of the TEF was a triumph of commitment and professionalism. ...The TEF has provided material for endless analyses, exploring the complexity of the datasets and the pattern of findings ... I am sure that institutions and assessors alike will be taking those insights forward into future rounds of the TEF.

Husbands began his foreword by saying how the TEF has built on the Research Assessment and Research Excellence exercises. This is an odd point to make as there is very little similarity or overlap except that both the RAE/REF and the TEF are based upon a belief that teaching and research can be assessed remotely, a belief not everyone shares (Deem 2016). The RAE/REF has always distributed money to those doing well, which the TEF does not. Also Husbands focused a good deal on the process, whereas, as spelt out in this article, one of the major issues relates to the content and type of metrics used, not the process of assessing them. The HEPI report, by contrast, focuses on lessons learned drawn from the submissions of a sample of 12 representative HEIs (Beech 2017).

What we also do not have, unlike the RAE/REF (Deem 2016), is any detail about the *cultural and social* processes of the TEF panels' work (Lamont 2009). There is no analysis of where the panel members come from, why they were chosen or what they did or thought about the exercise beyond Husbands (Beech 2017) stating that the process was well run. Nor do we know the precise relationship between the metrics and written self-assessment submissions. Gaming of TEF is as likely as gaming of REF (Lucas 2006) but it is not yet clear what forms TEF gaming can take. Some

gaming is evident in who entered TEF 3, since most were those trying to raise their grade. In future new metrics may be added to TEF, such as learning gain and teaching intensity (Ashwin 2017). But teaching intensity (hours of tuition) counts may be meaningless without evidence that contact hours are a useful proxy for teaching excellence. Furthermore, as the TEF has now become the Teaching Excellence Framework and Student Outcomes Exercise, it seems market competition for graduates paying back the loan fast (those in highly paid jobs) may presage a move away from teaching excellence, as the major determinants of getting such jobs are social class, gender and degree subject, not teaching quality (Social Mobility Commission 2017). However, the then Higher Education Minister, Chris Skidmore, as noted earlier, indicated that he wanted to reflect on questions of the economic and social value of different occupations and the relationship with pay levels in graduate jobs.

When we look at some characteristics of the TEF outcomes to date, it appears that Paradeise and Thoenig's typology does not work as well for TEF as it does for standard international rankings and reputational exercises (see "Appendix"). The HEPI report (Beech 2017), while praising aspects of the TEF process and discussing lessons learned from submissions, confirms that the results of TEF 2 upset the hierarchy of UK HEIs. We found that the 'Venerable' category is hard to apply to most UK HEIs. Instead we have devised a new category of 'small specialist institutions,' typically focused on arts or humanities, some of whom did very well in the TEF and often do well in the REF too but are not really comparable with comprehensive HEIs. Though some HEIs that might be seen as 'Top of the Pile' entered TEF (e.g. Oxford and Cambridge) and got Gold, there are also many 'Wannabees' and 'Missionaries' in the Gold category too. In the 2017 TEF 2, which was the most populated of TEF 2 and TEF 3, 18 Gold award winners out of 45 were not in the international rankings at all and 14 more were below the top 300. Twenty-six Gold award winners from 2017 and 2018 got a grade point average of less than 3 in the REF 2014 (3* is for internationally renowned work and 4* for world leading research and only 3* and 4* work is funded) and a further six did not enter the REF 2014. Only 10 out of 24 members of the Russell Group of elite universities got Gold, and two of those only did so on their second attempt. Though the 'Missionaries' tend to cluster in Silver and Bronze, in the latter there were also 'Wannabees' (e.g. Goldsmiths, SOAS, St George's) and one 'Venerable' (London School of Economics). So if the TEF was intended to upset the dominant hierarchy, it has certainly begun to do that. It may also suggest that we are still some way from determining what shapes teaching excellence and what drives access to graduate jobs.

Who gains from TEF and who loses?

It is probably self-evident that those who get Gold are the winners but with around 50 institutions getting this accolade it is not exactly a scarce award and as Ashwin (2018) noted, after the TEF 3 results were released, five institutions moved to a Gold rating from a lower grade in just 6 months. This raises questions about volatility

and whether a 3 year-old TEF Gold result will mean anything at all to HE applicants (Ashwin 2018). The plan to allow higher fees to be charged by Gold and Silver award holders is in abeyance because of the current investigation in England under the Augar review into the costs of HE to students. The TEF also has no mechanism for, or any intention of, rewarding teachers themselves. Of course the TEF is a competition with winners and losers but it also discourages collaboration on teaching between English HEIs. Those with Silver are working out how to game their way to Gold and those with Bronze how to get Silver. If this really did improve the degree results and life chances of students from black and ethnic minority groups, mature students, women (who do well in HE but badly in the labour market) part-time students or those with a disability, that would be highly beneficial, but the employment data are unlikely to show these groups getting highly paid jobs as that involves a whole world of labour market discrimination not within the control of HEIs. Those institutions that have achieved Bronze may be bad at teaching or have working-class students who do not enter well-paid jobs in the city or live in expensive halls of residence (e.g. the so called ‘London effect’). Getting Bronze does have the capacity in a crowded student market place to lose institutions potential students and drop places in other rankings. Bronze awards may also mean international agents and student sponsors look elsewhere, though it probably also depends on the institution’s social capital—this is less of an issue for the London School of Economics than a former polytechnic. In the early stages of the TEF it was even threatened that Bronze award holders would not be allowed to recruit international students, though ironically with Brexit it will probably be harder to recruit international students anyway. HE applicants already have a wealth of information available about different institutions and are in danger of being overwhelmed by all the data. So the real winners of the TEF are almost certainly not going to be students, even though many would like lower fees and more contact hours, as the YouthSight 2018 student experience survey from the Higher Education Policy Institute showed (Neves and Hillman 2018).

Unintended consequences of the TEF

Some of the unintended consequences of the TEF include using an array of somewhat arbitrary metrics, emphasising the competitive elements of the TEF at the expense of seeing excellent teaching as also arising from collaboration, suggesting to recent graduates that unless they get a very well-paid private sector job they are a failure, and focusing attention on the preparation for and game-playing in the TEF rather than on supporting those diverse HE staff (some part of the precariat workforce of English academe), who contribute to excellent teaching. In purporting to be a way in which students select their programme, the focus is not on interesting content or innovative programmes but on the ‘value for money’ element. The previous Minister for Universities, Science, Research and Innovation, Sam Gyimah called for technology companies to use published data on graduate earnings related to students who had studied different courses to create a ‘Money Supermarket’ of

HEIs (Gyimah 2018). Of course, class, ethnicity, gender, ‘race’ and disability can and do significantly affect education outcomes but degree outcomes also relate to students making an effort to ensure that they get value for money by turning up for classes, studying hard and completing assessments to the best of their ability. The TEF merely reinforces the view that HE is all about students purchasing a product which is effectively a private good, rather than something which involves significant student effort and that can contribute to the public good (Deem and McCowan 2018; Marginson 2018).

Reflections and conclusions

We have seen that the TEF is ideologically and somewhat arbitrarily driven in relation to its metrics and also does not follow most of the recommendations of the literature on assessing teaching excellence or high quality teaching. But then TEF is no longer largely about teaching excellence; it is yet another mechanism for constructing a market in higher education in the UK. The TEF has a number of unintended as well as intended consequences. The TEF 2 and TEF 3 rankings have not followed Paradise and Thoenig’s typology of HEIs and academic quality and at least four or five different types of institution are in the Gold category. The TEF also encourages competition over collaboration, yet more inter-institutional collaboration could be very beneficial for teaching excellence.

The TEF looks set to become part of the HE landscape in England for the foreseeable future, even though it scarcely measures teaching excellence and relies on remote judgments and convenient metrics, not visits to institutions to inspect and observe the teaching of those HEIs awarded Gold which Ashwin (2015) and others advocate. The Office for Students is also challenging established worldwide methods of HE quality assessment and enhancement in favour of allowing the market and metrics alone to shape both. As the TEF increases in complexity it may well drop the self-assessment. Academics will be performance-managed to improve their teaching but receive no rewards. Students will still be confused about how to choose their programme and HEI. Though the focus on split metrics around ethnicity and gender may lead to greater support for widening participation students, this is unlikely to also effect change in the labour market. Finally, the tide of unrest around Vice Chancellor pay, as well what English HEIs’ purposes actually are, following the 2018 University and College Union USS pensions strike which more than a few students supported, is leading to greater questioning of the validity of the market-driven approach.

An intelligent accountability system to improve teaching quality in HE would be based upon the extant literature and would continue to evaluate the effects of the metrics to ensure that teaching quality was enhanced and that other, unintended effects were managed. As outlined above, the TEF contains no measure, even of the most rudimentary kind, of the professional support or qualifications in the teaching of HE staff. Nor does it take into account the high percentage of teaching staff

on precarious contracts in many UK HEIs, including some of those awarded TEF Gold. It does not even consider curriculum design. There is no causal model underlying the metrics and notion of teaching excellence. Instead, an index of variables is assembled to capture data on a range of issues that have been raised about HE in England. As a management tool with which to regulate HE, the purpose of the TEF is not to produce excellence in teaching. ‘Teaching Excellence Framework’ is a misnomer, but it is too late for politicians to reconsider its name except for adding ‘Student Outcomes’. Indices are only useful to the extent that they coincide with broader understandings of an issue. When the retail price index fails to capture an important aspect of inflation, it is sometimes adapted. We do not propose that the TEF is merely accepted as an index and that all arguments relating to the need for a causal model should be dropped. Our purpose in drawing attention to this issue is to point out that policy makers, at least in this instance, have a different model of reality, one which is socially constructed.

The Office for Students will find it hard to justify TEF outcomes if they appear to show completely different university rankings than would have been anticipated by other quality markers, including reputation. To that extent, the use of narratives has helped to ameliorate early disrepute of the TEF, by bringing about HEI rankings that are closer to those that would have been expected, but the results are still unexpected in the array of types of HEI getting Gold. The use of the TEF to construct a market in HE in England is ideologically driven and comes at an unfortunate time, with Brexit (the fate of which was still uncertain at the time of writing). EU students may be less attracted to study in England following the UK’s exit from the EU (as except in Scotland fees will become very expensive), pressurising the viability of some institutions. If they are branded third class, ‘Bronze’, HEIs’ chances of competing in the entire international HE market are further compromised. Since the point of a market is to make a profit and the Office for Students is charged with introducing new suppliers to the market, the fate of less strong institutions would be viewed as of secondary importance in a neoliberal worldview. What counts as intelligent accountability very much depends upon the underlying value system, as this case demonstrates.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: 2017 and 2018 TEF Gold awards by THE World Rankings 2018, REF 2014 grade point average, Paradise and Thoenig (2015) categorisation, Russell Group membership

HEI name	TEF result 2018	THE World Rankings 2018	REF 2014 GPA	Year of award + previous award where relevant	Category of HEI on basis Paradise and Thoenig	Russell Group universities
Arts Univ Bournemouth	Gold	NR	2.30	2017	Small specialist institution	
Aberystwyth Univ	Gold	301–350	2.84	2018	Wannabee	
Aston Univ	Gold	351–400	3.05	2017	Wannabee	
Bangor Univ	Gold	301–350	2.99	2017	Wannabee	
Bath Univ	Gold	251–300	3.17	2017	ToP	
Birmingham Univ	Gold	= 141	3.07	2017	ToP	✓
Bishop Grosseteste Univ	Gold	NR	1.96	2017	Missionary	
Buckingham Univ (nfp)	Gold	NR	NE	2017	Private not for profit	
Cambridge Univ	Gold	2	3.33	2017	ToP	✓
Conservatoire Dance and Drama	Gold	NR	NE	2017	Small specialist institution	
Coventry Univ	Gold	601–800	2.67	2017	Missionary	
De Montfort Univ	Gold	601–800	2.67	2017	Missionary	
Derby Univ	Gold	801–1000	2.07	2017	Missionary	
Dundee Univ	Gold	187	3.03	2017	ToP	
Durham Univ	Gold	97	3.14	2018 Silver 2017	ToP	✓
East Anglia Univ	Gold	= 188	3.11		ToP	
Edge Hill Univ	Gold	NR	2.22		Missionary	
Essex Univ	Gold	251–300	3.05		Wannabee	
Exeter Univ	Gold	= 130	3.08		ToP	✓
Falmouth Univ	Gold	NR	2.04		Missionary	
Guildhall Sch Music/Drama	Gold	NR	2.44	2018 Silver 2017	Small specialist institution	
Harper Adams Univ	Gold	NR	2.66	2018	Small specialist institution	
Hertfordshire Univ	Gold	501–600	2.61	2018 Silver 2017	Missionary	

HEI name	TEF result 2018	THE World Rankings 2018	REF 2014 GPA	Year of award + previous award where relevant	Category of HEI on basis Paradise and Thoenig	Russell Group universities
Huddersfield Univ	Gold	601–800	2.65	2017	Missionary	
Imperial College	Gold	8	3.36	2017	ToP	✓
Kaplan Open Learning (Pv)	Gold	NR	NE	2018	Private for profit	
Keele Univ	Gold	501–600	2.88	2017	Wannabee	
Kent Univ	Gold	301–350	2.95	2017	Wannabee	
Lancaster Univ	Gold	= 150	3.15	2017	ToP	
University of Law (Pv)	Gold	NR	NE	2017	Private for profit	
Leeds Univ	Gold	139	3.13	2017	ToP	✓
Lincoln Univ	Gold	601–800	2.54	2017	Missionary	
Liverpool Hope Univ	Gold	NR	2.21	2017	Missionary	
Liverpool Institute Perf Arts	Gold	NR	NE	2017	Small specialist institution	
Loughborough Univ	Gold	301–350	2.95	2017	Wannabee	
Newcastle Univ	Gold	= 175	3.09	2017	ToP	✓
Northampton Univ	Gold	NR	2.09	2017	Missionary	
Norwich Univ of Arts	Gold	NR	2.69	2017	Missionary	
Nottingham Univ	Gold	= 147	3.09	2017	ToP	✓
Nottingham Trent Univ	Gold	601–800	2.59	2017	Missionary	
Oxford Univ	Gold	1	3.34	2017	ToP	✓
Portsmouth Univ	Gold	501–600	2.75	2017	Missionary	
Robert Gordon Univ	Gold	801–1000	2.35	2017	Missionary	
Rose Bruford College	Gold	NR	NE	2017	Small Specialist institution	
Royal Acad of Music	Gold	NR	2.75	2018	Small specialist institution	
Royal Central School Speech and Drama	Gold	NR	3.03	2017	Small specialist institution	

HEI name	TEF result 2018	THE World Rankings 2018	REF 2014 GPA	Year of award + previous award where relevant	Category of HEI on basis Paradise and Thoenig	Russell Group universities
Royal College of Music	Gold	NR	3.01	2017	Small specialist institution	
Royal Northern College of Music	Gold	NR	3.08	2017	Small specialist institution	
Royal Vet College	Gold	351–400	3.11	2017	Small specialist institution	
St Andrews Univ	Gold	= 143	3.13	2017	ToP	
Surrey Univ	Gold	251–300	2.98	2017	Wannabee	
Swansea Univ	Gold	251–300	3.09	2018	Wannabee	
Univ of West of England	Gold	601–800	2.70	2018 Silver 2017	Missionary	
Univ of York	Gold	= 137	3.17	2018 Silver 2017	ToP	✓

Key: Nfp—private not for profit, Py—private for profit, NR—not ranked in Times Higher 2018 league tables, NE—not entered for REF 2014, ToP—Top of the Pile

References

- Adams, R. (2019). EU students could face higher fees to study in UK from 2020. *The Guardian*. Retrieved October 17, 2019 from <https://www.theguardian.com/education/2019/apr/28/eu-students-could-face-higher-fees-to-study-in-uk-from-2020>.
- Ashwin, P. (2015). *Reflective teaching in higher education*. London: Bloomsbury.
- Ashwin, P. (2017). *What is the teaching excellence framework in the United Kingdom, and will it work?* Retrieved February 2018, from http://eprints.lancs.ac.uk/83744/2/_88_ASHWIN_TEF_UK_HBU_without_track_changes_Final.pdf.
- Ashwin, P. (2018). *TEF 2018: Can students trust the gold standard?*. London: Times Higher Education.
- Baird, J. (2007). Alternative conceptions of comparability. In P. E. Newton, J. Baird, H. Goldstein, H. Patrick, & P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards* (pp. 124–156). London: Qualifications and Curriculum Authority.
- Baird, J., Cresswell, M. J., & Newton, P. N. (2000). Would the real gold standard please step forward? *Research Papers in Education*, 15(2), 213–229.
- Barkas, L. A., Scott, J. M., Poppitt, N. J., & Smith, P. J. (2017). Tinker, tailor, policy-maker: Can the UK government's teaching excellence framework deliver its objectives? *Journal of Further and Higher Education*, 43, 801–813.
- Beech, D. (2017). *Going for Gold: Lessons from the TEF provider submissions*. Oxford: Oxford Higher Education Policy Institute.
- Behle, H., Atfield, G., Elias, P., Gambin, L., Green, A., Hogarth, T., et al. (2015). Reassessing the employment outcomes of higher education. In J. Case & J. Huisman (Eds.), *Researching higher education: International perspectives on theory, policy and practice* (pp. 114–131). New York: Routledge Press.
- Brown, P., & Lauder, H. (2012). *The global auction: The broken promises of education, jobs, and incomes*. Oxford: Oxford University Press.
- Cardoso, S., Rosa, M. J., & Stensaker, B. (2015). Why is quality in higher education not achieved? The view of academics. *Assessment and Evaluation in Higher Education*, 41(6), 950–965.
- Christie, T., & Forrest, G. M. (1981). *Defining public examination standards (Schools Council Research Studies)*. London: Macmillan Education.
- Coe, R. (1999). *Changes in examination grades over time: Is the same worth less?* Paper presented at the annual conference of the British Educational Research Association, Brighton.
- Coe, R. (2007). Common examinee methods. In P. E. Newton, J. Baird, H. Goldstein, H. Patrick, & P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards* (pp. 331–367). London: Qualifications and Curriculum Authority.
- Coe, R. (2010). Understanding comparability of examination standards. *Research Papers in Education*, 25, 271–284.
- Corbett, A. (2016). *Brexit was a huge shock for universities. Now we must regroup and deepen our European links*. Retrieved October 17, 2019 from <https://blogs.lse.ac.uk/brexit/2016/07/01/brexit-was-a-huge-shock-for-universities-now-we-must-regroup-and-deepen-our-european-links/>.
- Cresswell, M. J. (1996). Defining, setting and maintaining standards in curriculum-embedded examinations: Judgemental and statistical approaches. In H. Goldstein & T. Lewis (Eds.), *Assessment: Problems, developments and statistical issues* (pp. 57–84). Chichester: John Wiley.
- Deem, R. (2015). *A critical commentary on Ray Land and George Gordon 'Teaching excellence initiatives: Modalities and operational factors'*. New York: Higher Education Academy.
- Deem, R. (2016). Recent research evaluations in the UK and Portugal: Methodologies, processes, controversies and consequences. In C. Sarrico, P. Teixeira, & A. Magalhaes (Eds.), *Global challenges, national initiatives, and institutional responses—The transformation of higher education* (pp. 159–186). Rotterdam: Sense Publishers.
- Deem, R. (2017). New managerialism in higher education. In J. C. Shin & P. Texeira (Eds.), *Encyclopaedia of international higher education systems and institutions*. Dordrecht: Springer.
- Deem, R., & Brehony, K. J. (2005). Management as Ideology: The case of 'new managerialism' in higher education. *Oxford Review of Education*, 31(2), 213–231.
- Deem, R., & Magalhaes, A. (2017). *University governance reforms in European Higher Education; some reflections on the roles and interventions of external governing board stakeholders in Portugal and the UK*. Unpublished, European Higher Education Society, Universidade do Porto, Portugal.

- Deem, R., & McCowan, T. (2018). Understanding the role of graduates in society: Which conception of public good?, Chapter 5. In P. Ashwin & J. Case (Eds.), *Higher education pathways: South African undergraduate education and the public good* (pp. 61–78). Cape Town: African Minds.
- Department for Business, E. a. I. S. (2016). *Building on success and learning from experience an independent review of the research excellence framework (the stern report)*. London: Department of Business, Energy and Industrial Strategy.
- Department for Business Innovation and Skills. (2015). *Fulfilling our potential; Teaching excellence, social mobility and student choice (the higher education green paper)*. London: Her Majesty's Stationery Office.
- Department for Business Innovation and Skills. (2016). *Success as a knowledge economy: Teaching excellence, social mobility and student choice*. London: Department for Business Innovation and Skills.
- Department for Education. (2017). *Teaching excellence and student outcomes framework specification*. London. Retrieved October 17, 2019 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/658490/Teaching_Excellence_and_Student_Outcomes_Framework_Specification.pdf.
- Education Committee. (2012). *The administration of examinations for 15–19 year olds in England*. House of Commons. Retrieved June 12, from <https://publications.parliament.uk/pa/cm201213/cmsel ect/cmeduc/141/14102.htm>.
- Elton, L. (1998). Dimensions of excellence in university teaching. *International Journal for Academic Development*, 3(1), 3–11.
- European Commission High Level Group on the Modernisation of Higher Education. (2013). *Improving the quality of teaching and learning in Europe's higher education institutions*. Retrieved October 17, 2019 from http://ec.europa.eu/education/higher-education/doc/modernisation_en.pdf.
- Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. New York: Free Press.
- Gaus, N. (2015). *The Indonesian state university in flux: Academics and the neo-liberal turn*. PhD, Manchester unpublished.
- Gibbs, G., Knapper, C. K., & Piccinin, S. (2009). *Departmental leadership of teaching in research-intensive environments*. London: Leadership Foundation for Higher Education.
- Gyimah, S. (2018). We want students to be better informed about degree choices and the returns—Today, we're officially launching a competition for tech companies to take graduate data and create a MoneySuperMarket for students, giving them real power to make the right choice [Tweet].
- Hanson, F. A. (2000). How tests create what they are intended to measure. In A. Filer (Ed.), *Assessment: Social practice and social product* (pp. 67–81). London: Routledge.
- Hunter, F., & De Wit, H. (2016). Brexit and the European shape of things to come. *International Higher Education*, 87, 2–3.
- Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations? *Journal of Economic Education*, 36, 29–42.
- Johnson, J. (2017). *Embracing accountability and promoting value for money in higher education: Speech to UUK annual conference*. Retrieved August 13, 2018, from <https://www.gov.uk/government/speeches/jo-johnson-speech-to-uuk-annual-conference>.
- Krücken, G. (2014). Higher education reforms and unintended consequences: A research agenda. *Studies in Higher Education*, 39(8), 1439–1450.
- Lamont, M. (2009). *How professors think: Inside the curious world of academic judgment*. Harvard: Harvard University Press.
- Land, R., & Gordon, G. (2015). *Teaching excellence initiatives: Modalities and operational factors*. New York: Higher Education Academy.
- Langbein, L. (2008). Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review*, 27, 417–428.
- Lucas, L. (2006). *The research game in academic life*. Maidenhead: Open University Press and the Society for Research into Higher Education.
- MacNell, L., Driscoll, A., & Hunt, A. N. (2014). What's in a name? Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, 49, 291–303.
- Marginson, S. (2018). Public/private in higher education: A synthesis of economic and political approaches. *Studies in Higher Education*, 4(2), 322–337. <https://doi.org/10.1080/03075079.2016.1168797>.
- Mayhew, K. (2017). UK higher education and Brexit. *Oxford Review of Economic Policy*, 33, 155–161.

- McKie, A. (2018). Hertfordshire's £66K fine is first major OfS sanction. *Times Higher Education*. London: THE.
- McPherson, M. A. (2006). Determinants of how students evaluate teachers. *Journal of Economic Education*, 37(1), 3–20.
- McPherson, M. A., & Todd Jewell, R. (2007). Leveling the playing field: Should student evaluation scores be adjusted? *Social Science Quarterly*, 88(3), 868–881.
- Merton, R. K. (1936). The unanticipated consequences of purposive social action. *American Sociological Review*, 1(6), 894–904.
- Milburn, A. (2012). *University challenge: How higher education can advance social mobility: A progress report by the independent reviewer on social mobility and child poverty (the Milburn report)*. Retrieved August 2013, from <http://socialwelfare.bl.uk/subject-areas/services-activity/education-skills/cabinetoffice/137068Higher-Education.pdf>.
- Neves, J., & Hillman, N. (2018). *Student academic experience survey*. Oxford: Higher Education Policy Institute and Advance HE.
- Newton, P. E. (1997a). Examining standards over time. *Research Papers in Education*, 12(3), 227–248.
- Newton, P. E. (1997b). Measuring comparability of standards between subjects: Why our statistical techniques do not make the grade. *British Educational Research Journal*, 23, 433–449.
- Newton, P. E. (2003). *Contrasting definitions of comparability*. Paper presented at the QCA standards and comparability seminar, April. Milton Keynes.
- Newton, P. E. (2005). Examination standards and the limits of linking. *Assessment in Education: Principles, Policy and Practice*, 12(2), 105–123.
- Newton, P. E. (2010). Contrasting conceptions of comparability. *Research Papers in Education*, 25(3), 285–292.
- O'Neill, O. (2002). *A question of trust: The BBC Reith lectures 2002*. Cambridge: Cambridge University Press.
- O'Neill, O. (2013). Intelligent accountability in education. *Oxford Review of Education*, 39(1), 4–16.
- Ofqual. (2014). *Setting GCSE, AS and A level grade standards in summer 2014 and 2015*. Ofqual/15/5759. Retrieved October 17, 2019 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/451321/2015-08-05-summer-series-gcse-as-and-a-level-grade-standards.pdf.
- Paradeise, C., & Thoenig, J.-C. (2015). *In search of academic quality*. London: Palgrave Macmillan.
- Perkins, G. (2018). The teaching excellence framework (TEF) and its impact on academic identity within a research-intensive university. *Higher Education Policy*, 32, 297–319.
- Power, M. (1997). *The audit society*. Oxford: Oxford University Press.
- Robinson, W., & Hilli, A. (2016). The English teaching excellence framework and professionalising teaching and learning in research-intensive universities: An exploration of opportunities, challenges, rewards and values from a recent empirical study. *Foro de Educación*, 14(21), 151–165.
- Ross, J. (2019). Performance funding stirs fears of unfair penalties. *Times Higher*. 15th August (p. 7).
- Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change*, 11, 45–61.
- Sahlberg, P. (2013). The most wanted: Teachers and teacher education in Finland, chapter 1. In L. Darling-Hammond & A. Lieberman (Eds.), *Teacher education around the world: Changing policies and practices* (pp. 1–21). London: Taylor and Francis. <https://doi.org/10.4324/9780203817551>.
- Savage, M. (2019). Universities to be fined for awarding too many top degrees. *Times Higher Education*. London: THE.
- Social Mobility Commission. (2017). *State of the nation 2017: Social mobility in Great Britain*. Retrieved October 17, 2019 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/662744/State_of_the_Nation_2017_-_Social_Mobility_in_Great_Britain.pdf.
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching. *Review of Educational Research*, 83(4), 598–642.
- Stroebe, W. (2016). Why good teaching evaluations may reward bad teaching: On grade inflation and other unintended consequences of student evaluations. *Perspectives on Psychological Science*, 11(6), 800–816.
- Subtirelu, N. C. (2015). “She does have an accent but...”: Race and language ideology in students’ evaluations of mathematics instructors on RateMyProfessors.com. *Language in Society*, 44(1), 35–62.
- Tesio, L. (2014). Causing and being caused: Items in a questionnaire may play a different role, depending on the complexity of the variable. *Rasch Measurement Transactions*, 28, 1454–1456.

- Tsui, A. B. M. (2015). *A critical commentary on Ray Land and George Gordon 'Teaching excellence initiatives: Modalities and operational factors'*. New York: Higher Education Academy.
- Universities UK. (2017). *Review of the teaching excellence framework: Year 2*. London: Universities UK. Retrieved October 17, 2019 from <http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2017/review-of-the-teaching-excellence-framework-year-2.pdf>.
- Veiga, A., Magalhães, A., & Amaral, A. (2015). From collegial governance to Boardism: Reconfiguring governance in higher education. In J. Huisman, H. De Boer, D. Dill, & M. Souto-Otero (Eds.), *The Palgrave international handbook of higher education policy and governance* (pp. 398–416). London: Palgrave.
- Watermeyer, R. (2012). Issues in the articulation of 'impact': The responses of UK academics to 'impact' as a new measure of research assessment. *Studies in Higher Education*, 39(2), 359–377.
- Wilsdon, J. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. Bristol: Higher Education Funding Council for England.
- Wood, M., & Su, F. (2017). What makes an excellent lecturer? Academics' perspectives on the discourse of teaching excellence in higher education. *Teaching in Higher Education*, 22(4), 451–466.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.