

## Russell Group response to the independent review of TEF call for views

### 1. Summary

- 1.1 We do not believe subject-level TEF as currently conceived delivers sufficient value to prospective applicants and institutions. There is a high risk it will produce misleading results given the statistical flaws inherent in the proposed model, even if this can be moderated to some extent with contextual statements. In turn, this means subject-level TEF is likely to mislead prospective students.
- 1.1 **Subject-level TEF should not therefore be taken forward.** Instead, efforts should be made to support prospective students to use existing information more effectively through the creation of a new information interface.
- 1.2 **We also recommend making several key reforms to provider-level TEF** to address significant flaws in the benchmarking methodology and to make the exercise more fit for purpose in supporting prospective applicants in their decision-making. These include:
- placing substantially more weight on absolute values alongside benchmark and sector average scores
  - removing the flawed 'grade inflation' metric and ensuring institutions subject to NSS boycotts, or action which could skew NSS results, are not penalised
  - replacing the blunt medal rating system with a "profile approach" which could involve providing much more helpful information about institutional strengths and weaknesses on a dashboard where informed comparisons can be made.
- 1.3 Any further development of TEF at provider-level should be undertaken in collaboration with institutions in England, and in devolved administrations if they wish to participate, as well as academic staff to build buy-in for the exercise from the academic community. A clear strategy is also needed to mitigate the risk that TEF outcomes could damage the reputation of UK HE overseas. In considering the future model of TEF, we must also reflect on the negative impact the framework is having on existing tools to support enhancement, including the NSS.

### 2. Context

- 2.1 We welcome the opportunity to submit evidence to the independent review of TEF through the call for views process. We particularly welcome the focus on the objectives of TEF and whether it is fit for purpose. In order to add genuine value for prospective applicants and institutions, any TEF exercise should meet the following criteria:
- results should be methodologically and statistically sound
  - information should be sufficiently granular and comprehensive
  - enhancement and innovation in new types of provision should be incentivised
  - costs should be proportionate to the value of the exercise, especially as these will be met at least partially from students' fees and public funding.

### 3. Issues with the benchmarking methodology

- 3.1 We are concerned the TEF methodology is not statistically sound and the outcomes it produces will therefore risk misleading prospective applicants. There are a number of reforms which could be made to the TEF framework to improve the validity of the results and

encourage a focus on enhancement, but it will be very difficult to eradicate all statistical flaws, especially at subject-level.

*High absolute performance is not sufficiently recognised*

- 3.2 High absolute performance is not sufficiently recognised and rewarded. Universities that have consistently maintained very high performance may not be able to deliver continuous and significant positive differences from benchmark values, and as performance on the metrics improves, the 2% substantive difference threshold test (materiality test) becomes harder, and sometimes practically impossible, to achieve.
- 3.3 Whilst the Government has introduced a mechanism to identify the top 10% of provider scores for each metric, these absolute value markers do not directly contribute to the calculation of the initial hypothesis – and therefore to the final TEF award.
- 3.4 Very high performance can still go unrecognised, for example in two instances where applicants could be seriously misled:
- data published for the TEF4 exercise shows six HEIs including two Russell Group universities achieved continuation scores of over 95% (i.e. fewer than 1 in 20 students did not continue) but were not recognised with either a flag or an absolute value marker. Meanwhile, another provider achieved a positive flag with a continuation rate of just over 70% (i.e. three in every ten students dropping out).
  - similarly, the TEF4 data shows over 96% of students at one Russell Group institution progressed into employment or further study but this score did not feature within the top 10% of the sector and the university's performance was not recognised through the flag system as it did not beat the benchmark score.
- 3.5 Whilst the assessment of the TEF metrics tells us about how well institutions perform against expectations for their particular student intake, research has found no evidence that students understand that the outcomes of the process are based on performance against a benchmark rather than in absolute terms.<sup>1</sup> This issue is further compounded by the fact that the TEF differs from many other sources of information which students will be familiar with such as KIS and league tables which do focus on absolute values and are updated annually.

**Recommendation 1:**

- 3.6 **To recognise and incentivise high performance more effectively and to reflect prospective students' understanding of TEF ratings, absolute values should be used to generate the initial hypothesis at provider level.** An appropriately weighted combination of actual values and significant differences from benchmark scores could be used to assess performance. For example, for institutions with a high benchmark for a metric, increased weight could be given for the absolute value score achieved against that metric to recognise that the benchmark will be very difficult to beat.

*Universities contributing to the generation of their own benchmarks makes them harder to beat*

- 3.7 We have raised concerns about underlying problems in how the benchmarks are created. **Universities contribute to their own benchmarks, which in turn makes them harder to beat especially for larger providers.** In the extreme case, a university that makes 100%

---

<sup>1</sup> 'Teaching excellence: the student perspective' Trendence research commissioned by a consortium of students' unions (2017)

contribution to its benchmark could only ever achieve its benchmark value. The following example using the TEF4 data illustrates the point:

- the continuation metric, when ranked by contribution to benchmark, shows that only 13 significance flags were achieved by those HEIs with an above average contribution to benchmark and 18 flags were achieved by those HEIs with a below average contribution – so institutions with a below average contribution to their benchmark are almost 40% more likely to achieve a flag compared to those with an above average contribution.

3.8 This issue could be mitigated in part if the methodology is amended to place more focus on absolute values in the generation of the initial hypothesis. We would encourage the Office for National Statistics to also consider this issue as part of their analysis for the review.

#### *The generation of significance flags risks disincentivising enhancement*

3.9 The generation of the initial hypothesis does not take account of the degree to which the significance flags have either been met or not been met, so patchy extreme performance may be rewarded over consistent strength. For example:

- in data published for the TEF4 exercise, one institution outperformed their benchmark on highly skilled employment and further study by 26 percentage points (with 94% of students entering highly skilled employment or further study after six months) and another institution outperformed their benchmark on the same metric by only 2 percentage points (with a score of just 73% of students progressing to highly skilled employment and further study). However, each provider received the same level of recognition with the outcome of a positive flag.

3.10 Rather than encouraging institutions to focus on the areas of performance which require the most improvement, the methodology creates perverse incentives to focus on areas where institutions are almost meeting or are slightly above their benchmark. **The way the metrics are assessed does not therefore sufficiently incentivise enhancement.**

#### *Small student numbers risk skewing outcomes*

3.11 We are also concerned that small student numbers could risk skewing the benchmarked assessment of the metrics, and this is a particular problem at subject-level and in the split metrics. **Even when students are aggregated into broad subject groups the data is based on such small numbers there is a risk that outcomes could be determined by random year-on-year fluctuations as opposed to genuine variations in quality. Aggregating subjects also undermines the usefulness of the data for prospective students who wish to use TEF to evaluate a particular course before application.**

3.12 Although we do not have access to any subject-level data to model, we believe that the smaller the denominator size the more difficult it will be to achieve a flag (negative or positive) and there will be a tendency to a neutral (=) score. We can test this (albeit crudely) with institution-level data for the core metrics and it shows that where the cohort size is less than 100, around 80% of the metrics were neutral (=). In contrast, where the cohort size was greater than 1,000 around 60% of the metrics were neutral. At subject-level and in the splits, this could result in data that is of very little use.

3.13 One way around this issue could be to introduce a sufficient cohort size threshold, although this would inevitably limit the scope of the exercise to institutions and subjects which are large enough to meet the threshold. While we can expect the impact of this to be minimal at the provider level, it is likely to cause extensive problems with data coverage at the subject level resulting in some subjects at some providers being excluded from participating in TEF. However this is explained in caveats to any subject-level TEF announcement, the perception

will linger that these subjects at these providers are not “up to scratch”, having been excluded from TEF; the associated damage would be significant.

- 3.14 Based on the methodology being piloted at subject-level, 87% of providers already have non-reportable core metrics in at least one subject, and for 27 subject areas at least 20% of providers will have non-reportable core metric(s). Introducing a sufficient cohort threshold would exacerbate this issue and so, in turn, limit the ability of prospective applicants to use TEF results as part of their decision-making.

#### *Maximising coverage at the expense of consistency*

- 3.15 As part of the subject-level pilots currently being conducted, where metrics are non-reportable due to small numbers, provider-level metrics are being substituted for the purposes of assessment. This means while some providers will receive an award based on assessment of a full suite of subject-level data, others will receive an award based on partial data with provider-level metrics used to plug any gaps.
- 3.16 **We are concerned that the coverage of TEF awards at subject-level has been prioritised at the expense of consistency in the way the awards are assessed.** This risks misleading prospective applicants who will not be able to distinguish between outcomes which have been awarded based on very different sets of data.

#### **4. Issues within the metrics**

- 4.1 The current metrics do not measure teaching quality directly, although this is one of the most important factors for students.<sup>2</sup> The peer review element of TEF is therefore critically important in enabling institutions to provide a wider range of qualitative and quantitative evidence regarding their approach to teaching and learning.
- 4.2 Student disengagement with the NSS has emerged as a long-term issue following its use in TEF. This has had significant knock-on consequences for universities who interrogate the results of the NSS to inform improvements in curriculum, teaching and learning quality, learning resources and academic support. This means, in turn, that the value of the survey in informing the enhancement of teaching and learning, and the broader student experience, has been diminished. **Indeed, it appears that TEF has itself damaged the NSS.**
- 4.3 There are pressing concerns about the validity of the NSS results for some institutions. Some universities have no published NSS metrics at all for TEF4, making it very difficult or in some cases impossible to achieve a gold initial hypothesis. The continued NSS boycott will therefore start to skew the results, disadvantaging those who have been affected. Other external events (such as the recent strike action) may also have compromised the validity of some NSS scores.<sup>3</sup> If these issues continue, it will become more difficult to continue using NSS scores in a robust way to inform TEF outcomes.

---

<sup>2</sup> Recent research on value-for-money by the Office for Students found students believe quality of teaching to be the single most important factor demonstrating that an institution offers good value for money.

<https://www.officeforstudents.org.uk/new-research-shines-spotlight-on-student-perceptions-of-value-for-money/>

<sup>3</sup> Anecdotally, some institutions have found in analysis of free text responses to the 2018 NSS that the number of instances of references to the strike action correlate with lower than expected scores. A comprehensive analysis of this issue should be undertaken by the Office for Students.

### Recommendation 2:

4.4 Whilst it may be helpful to consider other ways to capture student feedback, any new metric could encounter similar issues. In light of this, we recommend that the effectiveness of the NSS in capturing student satisfaction is monitored. In the meantime, **the TEF methodology should be amended to ensure institutions subject to NSS boycotts, or action which could skew NSS results, are not penalised as a result.**

4.5 More broadly, our concern remains that institutions are being assessed on their performance against metrics over which they do not have direct control. In addition, we have **a number of concerns about statistical anomalies in the metrics which could impair the validity of TEF results.** These include the following:

- The majority of the core metrics are based on sampling and it is important to acknowledge there could be a non-response bias impacting on the NSS and DLHE scores. The ONS has highlighted that neither the DLHE nor the NSS capture undergraduates who do not complete their courses and this is also the case for LEO data.
- The ONS also found comparisons between institutions using raw NSS data are not usually significant, meaning most university performances are statistically indistinguishable.<sup>4</sup> If more precise comparisons are made – e.g. by social class of students, or by subject of study – confidence intervals become much wider, meaning results are less accurate.
- DLHE and LEO do not take any account of regional difference in employment outcomes despite evidence students who study in their home area are likely to remain there for work after graduation.
- The inclusion of the grade inflation metric within the provider-level assessment fails to recognise that there are a number of factors which can affect trends in degree attainment including prior attainment, subject mix, student characteristics, improvements in teaching practice and student engagement, and so on.

### Recommendation 3:

4.6 Whilst the current proposal not to use a “grade inflation” metric at subject-level is welcome, **it is a flawed measure and should also be removed from provider level TEF.** If TEF is effective in improving teaching quality, we would expect attainment to rise over time. Trends in degree attainment represent a complex picture and make it difficult to isolate cause and effect.

## 5. Issues with the outcomes

5.1 The TEF descriptors (gold/silver/bronze) are blunt instruments which do not recognise the rigorous quality assurance system which institutions must satisfy in order to operate. They also create harsh cliff edges between bandings and there is a clear risk they will be taken at face-value and interpreted simplistically. Indeed, in their recent research DfE found applicants consider a bronze award to indicate ‘under achievement’, rather than an achievement beyond satisfactory practice.<sup>5</sup> The ranking system also has the consequence of labelling most of UK higher education provision as silver which is likely to be regarded as

<sup>4</sup> ‘Teaching Excellence Framework: Review of Data Sources’ Office for National Statistics (June 2016)

<sup>5</sup> ‘Evaluation of Provider-level TEF 2016-17 (Year 2): Measuring the initial impact of the TEF on the Higher Education landscape’ (January 2019)

'average'. This presents an inaccurate picture of quality across the sector, both in a domestic and international context.<sup>6</sup>

- 5.2 Research has found students consistently assume that a gold rating at one provider is directly comparable to a gold rating at another institution, rather than recognising the benchmarking methodology means they are not comparing like with like.<sup>7</sup> The same research also found some evidence that students may not interpret TEF ratings as intended, with negative consequences for social mobility: a minority (6%) of students would reconsider applying to, or not have applied to, their current institution if it had been rated gold, and this number is higher for BME students (10%).

#### **Recommendation 4:**

- 5.3 **We recommend the gold/silver/bronze descriptors be replaced by alternative options which better reflect the nature of the assessment exercise and recognise that all eligible institutions have already met a high-quality bar to participate. This would help to ensure that TEF outcomes are not misinterpreted internationally.**

- 5.4 The existing descriptors also fail to recognise the diversity of provision across the sector. It is vital that we recognise students are not a homogenous group. They do not all have the same needs and priorities and are likely to be motivated to enter higher education for very different reasons. Moving away from the current medal system which produces a single rating could help to reflect this.

#### **Recommendation 5:**

- 5.5 **We propose that providers are awarded with a profile rather than a single rating across a range of areas covered by the current metrics.** These could include, for example, teaching quality, learning environment, and student outcomes – and potentially highlight good practice in other areas outlined in written submissions such as learning resources, research-led teaching, or innovation. This would empower prospective students to identify areas of provision which are of most importance to them and use TEF in a more sophisticated way in decision-making. It would also better recognise the diversity of the sector and reflect strengths even where on aggregate a gold outcome would not have been awarded – and so incentivise enhancement.
- 5.6 One option could include moving away from a summative assessment altogether and presenting information on performance (both absolute and benchmarked<sup>8</sup>) compared to the sector average. This would enable prospective applicants to access much more granular information about institutional strengths and weaknesses on a dashboard where informed comparisons can be made. It would also remove the perverse incentive for institutions to focus more effort on areas where they are almost meeting or are slightly above their benchmark, rather than on making improvements across the board. We would encourage the review team to consider whether parallels could be drawn with the approach to presenting institutional performance under consideration as part of the development of the Knowledge Exchange Framework (KEF).<sup>9</sup> Institutional submissions would also need to be presented alongside published data. Similarly, lessons can be learned from the Research Excellence Framework (REF) assessments where outputs, impact and research environment are presented as a profile.

<sup>6</sup> UK universities are regularly ranked amongst the highest in the world across a range of league tables. For example, in the QS 2019 subject rankings UK universities came top in the world for a total of 13 subjects.

<sup>7</sup> See footnote 1

<sup>8</sup> Although noting our previous caveats about flaws in benchmarking that will need to be addressed

<sup>9</sup> See here for more information: <https://re.ukri.org/knowledge-exchange/knowledge-exchange-framework/>

## 6. Additional concerns on subject-level TEF

### *Accounting for interdisciplinary provision*

- 6.1 The subject-level TEF methodology has been designed predominantly for students wishing to study single honours courses. However, this does not reflect the way in which universities structure their courses, nor the student experience which is increasingly sought. The framework being piloted is problematic as it will not capture the unique experience of students who are studying joint or multiple honours programmes, those undertaking modules in other departments and other faculties, or those studying courses which cross discipline boundaries (for example, natural sciences and liberal arts degrees). Predictions about the future needs of the labour market suggest that there will be a growth in teaching, learning and research that is increasingly multi-and inter-disciplinary in nature.
- 6.2 The proposal to test an approach where students are counted pro rata in the subject-level metrics against each subject to which their course is mapped will mean some applicants will need to look at two or more subject-level TEF awards to try to gain an understanding of the experience they can expect to receive. DfE student research shows that among those wishing to study multi-programme courses – such as politics, philosophy and economics, or English and theatre – fewer than 3 in 10 prospective students could identify the subject groupings relevant to them in the proposed subject-level TEF methodology.<sup>10</sup> If students cannot identify the results which are relevant to them, they will not be able to utilise the outcomes of the subject-level TEF.
- 6.3 The proposed approach to accounting for courses which cannot be mapped into separate subject categories – i.e. using three broad subject groups instead – is not sophisticated enough to capture and recognise the unique nature of provision in this area effectively. In the long-run, this may affect the demand for such courses.
- 6.4 DfE's analysis of the proportion of subjects this would affect (2%) is based on HESA returns from providers who entered TEF in 2016/17, but it is unlikely to be accurate as there is no universal approach towards categorising provision of this nature and the data excludes institutions in Scotland where inter-disciplinary provision is more common. Indeed, at one Russell Group university over 90% of courses have an element of inter-disciplinary flexibility, and many have plans to grow this kind of provision significantly over the next five years to meet future labour needs.
- 6.5 It is difficult to see how any meaningful comparisons can be made between inter-disciplinary subjects at different institutions as part of an assessment framework which is entirely focused on providing information about quality in individual disciplines. Yet, if this issue cannot be addressed, institutions could be discouraged from offering inter-disciplinary programmes. In effect, **there is a danger subject-level TEF could cement the parts of the UK higher education that are subject to the TEF in the past rather than incentivising innovation in new areas of provision.**

### *Establishing appropriate subject groupings*

- 6.6 Whilst numbers of students on individual courses are too small to enable any meaningful analysis, aggregating students into large enough groups to enable analysis undermines the validity of the assessment exercise at subject-level. This will lead to false comparisons being made between relatively broad subject groups at different institutions, and so risk misleading prospective applicants rather than providing them with useful information.

---

<sup>10</sup> <https://www.gov.uk/government/publications/teaching-excellence-framework-and-informing-student-choice>

- 6.7 Crucially, there does not appear to be any evidence that teaching quality across different courses in the same groups under the Common Aggregation Hierarchy Level 2 (CAH2) will be similar. The Royal Statistical Society has highlighted that as the subject groups contain different numbers and mixtures of individual subjects and often in different departments or even faculties from institution to institution, there will be a lack of comparability between the awards. They conclude that this could “invalidate fair assessment” and undermine the intention of the exercise – to enable prospective students to make comparisons.<sup>11</sup> The DfE student research also found applicants struggled to identify the courses they are interested in studying even where these were “simple” courses (where courses are classified into one subject area).<sup>12</sup>
- 6.8 **Whilst a more granular version of the CAH2 system could be created to address some of the issues outlined above, this would exacerbate existing issues with small student numbers skewing the results and would increase the burden of the exercise as a whole.** Under the subject-level TEF model being piloted in 2018/19, participating institutions will be required to produce up to 193 pages worth of submissions if they are to be assessed across 35 subject areas. Introducing more granular subject categorisation would mean comprehensive universities would be required to produce even longer written submissions.

*‘Clustering’ of results in some subjects makes it difficult to produce meaningful comparisons*

- 6.9 For some subjects, the metrics are clustered together within a very small range, often but not always at the top end of the spectrum. This makes it **very difficult to produce meaningful comparisons** between them in terms of performance. For example, the performance of graduates in entering highly skilled employment and further study for medicine and dentistry and nursing is extremely high. This effect is likely to lead to a preponderance of golds in some subjects and, potentially, of bronzes in others. It is difficult to see how the outcomes of this process would offer applicants a useful source of information which enables them to differentiate between providers.
- 6.10 In cases where the natural distribution of ratings results in the majority of providers receiving bronze, this could create real incentives to close such courses if student demand drops off. This would impact on the diversity of provision, reduce student choice and is likely to have an impact on skills provision and ability for the sector to contribute towards delivering on the Industrial Strategy.

## **7. Impact on international competitiveness**

- 7.1 There is a significant risk that TEF outcomes could damage the reputation of UK HE overseas and so affect international student recruitment. This could threaten the financial sustainability of a number of strategically important courses, which may then have knock-on effects for domestic students and the UK economy more broadly.
- 7.2 A recent survey by QS found international applicants across a range of markets thought a gold rating in TEF was more important than a high ranking in a league table: 62% of applicants preferred a gold rating over a university ranked 25th or higher in the UK, compared to 50% for a silver rating and 37% for a bronze rating.<sup>13</sup> This demonstrates the importance of ensuring the approach is well understood across international markets.

<sup>11</sup> Royal Statistical Society: Response to the Teaching Excellence and Student Outcomes Framework, Subject-level Consultation (2018)

<sup>12</sup> <https://www.gov.uk/government/publications/teaching-excellence-framework-and-informing-student-choice>

<sup>13</sup> <https://www.internationalstudentsurvey.com/international-student-survey-2018/>

- 7.3 As with other students, it is unlikely international students will be aware that any institution which receives a “bronze” rating has already met a very high bar for quality through the existing QA system. Rather, this is likely to be seen as a mark of poor quality and could deter international student applications. Over 60% of respondents to the QS survey felt TEF had not been well explained to them and it is difficult to envisage how the complex benchmarking process could be sufficiently explained to international, as well as domestic applicants.
- 7.4 In addition to impact on individual students’ choices, a number of overseas governments also use measures of quality and league table rankings to make decisions, including: allocation of scholarships for overseas study; to shape national frameworks for cooperation; and to formulate transnational education (TNE) agreements based on league table performance. It is therefore possible TEF (especially at subject-level) could be used in these decisions in future, meaning that universities with a bronze, or even silver, rating could miss out to competitors in other countries, with a net negative impact for the competitiveness of UK HE.

#### Recommendation 6:

- 7.5 The Government has engaged with the sector on a script which is intended to assist with the way in which TEF results are understood internationally. Whilst transparency about how TEF works and what it measures is helpful, this will be insufficient to mitigate the damage which the TEF could have on perceptions of quality across UK HE. **A clear strategy is needed to mitigate any unintended negative consequences. This should form part of any strategy to promote the quality of UK HE overseas.**

## 8. Burden and value for money

- 8.1 We have significant concerns about the burden associated with delivering TEF, especially at subject-level. UUK has estimated that HEIs spent £4.1 million on staff costs to participate in the TEF2 provider-level exercise, and it is unclear to what extent this has had a positive effect on enhancement of teaching quality within institutions. A survey conducted by UUK found that the TEF has only influenced around 20% of responding institutions to review and revise existing learning and teaching enhancement activities, whilst much of the additional investment in this area had been planned before the TEF was introduced.<sup>14</sup>
- 8.2 As noted above, there is growing evidence prospective students do not understand what TEF is measuring and will not therefore be able to use the results to make meaningful judgements. In their recent study among prospective students, DfE found 66% of respondents believed the TEF assessment process involved inspectors observing teaching and speaking to students, rather than being based on proxy metrics and submissions.
- 8.3 Any meaningful and robust exercise at subject-level is likely to require a significant bureaucratic structure to support it, so we can expect this will lead to the costs at provider-level being multiplied substantially in a move to subject-level assessments. Following a provider cost survey, DfE estimates the cost of the subject-level TEF model currently being piloted to be at least £24 million<sup>15</sup> if applied to all UK providers, however this seems likely to be a significant under-estimate and indeed a recent analysis by UUK has estimated the cost to total £37.6 million.<sup>16</sup> There is an inevitable trade-off between the administrative burden which a subject-level TEF exercise will introduce and the extent to which such an exercise can deliver robust and meaningful results. Any changes to improve the rigour of the

<sup>14</sup> ‘Review of the Teaching Excellence Framework Year 2: Process, Results and Next Steps’ (2017)

<sup>15</sup> ‘Teaching Excellence and Student Outcomes Framework: Findings from the first subject pilot 2017-18’ Annex D: Provider cost survey (2018)

<sup>16</sup> <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2019/future-of-the-tef-independent-reviewer.pdf>

methodology (for example, more granular subject groupings, or placing more weight on institutional submissions over flawed data) would lead to an escalation in costs.

- 8.4 The Government intends to run assessment cycles every five years for subject-level TEF in order to limit the cost and burden of the exercise, which will lead to circumstances where applicants are utilising information based on data that is, in the case of LEO, up to 10 years old. The less frequently the assessments take place the less useful the results will be to prospective applicants and the less representative the data will be of a provider's current performance. Conversely, more frequent assessments would lead to even higher costs and the time-lag associated with some metrics data would mean assessments would still be based on outcomes for students who had graduated from institutions years previously.
- 8.5 We also encourage the review team to consider the broader context within which the sector will be operating, which further heightens the challenges associated with delivering TEF, especially at subject level. Any reduction in funding for teaching as a result of the post-18 review, coupled with the implications of Brexit and other financial challenges could undermine the financial sustainability of many universities and will make it even harder to deliver additional burdensome regulatory requirements. **In straitened circumstances, requiring universities to undertake an assessment process of this magnitude will inevitably mean resources will be diverted from core missions.**
- 8.6 Indeed, we need to consider the implications of rolling out the subject-level TEF in the same timeframe as the next iteration of the REF. Whilst the proposed deadline for submissions to subject-level TEF is early 2020, the costs associated with delivering REF are concentrated in the 18-24 months ahead of the submission deadline (scheduled for the end of 2020 for the forthcoming REF round). If these two frameworks were to coincide, this would place a considerable burden on academic and administrative teams within institutions and present logistical challenges to the resourcing of both the REF and TEF panels. A report on REF 2014 found for some institutions that "the burden of REF falls most heavily on senior academics and this in turn can reduce the institution's capacity to pursue other sources of research funding".<sup>17</sup> Multiplying the burden on academics and managers by aligning the REF and subject-level TEF timescales would be very likely to exacerbate this issue, squeezing out time for academics to undertake their usual work including research and teaching.

#### **Recommendation 7:**

- 8.7 The value which the TEF delivers to prospective applicants, institutions and the wider public must be proportionate to the cost, especially as this will be met from students' fees and public funding. **We do not believe subject-level TEF as currently conceived delivers sufficient value to prospective applicants and institutions.** In fact, there is a high risk it will produce misleading results given the flaws inherent in the proposed model (as outlined above) and will prove to be a distraction for institutions where work to enhance the quality of provision is already underway. An even greater concern is that subject-level TEF creates a high risk of misleading prospective students. **Subject-level TEF should not therefore be taken forward.**

---

<sup>17</sup> 'REF Accountability Review: Costs, benefits and burden' Report by Technopolis to the four UK higher education funding bodies (2015)

## 9. Proposed alternatives to subject-level TEF assessments

### *Information for prospective applicants*

- 9.1 Evidence suggests that rather than producing more information for prospective applicants, it would be better to support them to make more effective use of existing information. The OFT raised this issue in 2014, calling for better integration and clearer links between the various information sources already available.<sup>18</sup> Research from the AdvanceHE (formerly HEA) found that even where information has been readily available it has not been well used and needs to be accompanied by guidance.<sup>19</sup> Early evaluations of the Unistats website by HEFCE found a number of participants expressed concerns about being overwhelmed by data. This suggests that it is not necessarily additional data that is required, but an effective mechanism for filtering and making sense of existing data.
- 9.2 A new information interface is needed and should be developed according to the following operational characteristics:
- (a) **User-friendly and accessible platform:** the interface should support users to navigate through all information sources provided; it should be linked to by a range of stakeholders and information providers in order to raise awareness among potential applicants.
  - (b) **A site which collates and synthesises information from a wide range of different sources:** it should be as comprehensive as possible to tackle the problem of dispersed and (somewhat) incompatible information resources currently available. It should provide effective signposting to university websites.
  - (c) **Incorporate an element of guidance:** a new interface should support users to identify and find the relevant information for them by providing easy-to-understand guidance about what is available and how to find it.
  - (d) **Sophisticated enough to allow applicants to select information according to their own priorities, preferences and abilities:** this approach would enable users to filter and prioritise information to determine which institution and course best meets their needs.
  - (e) **Provide end to end support through the decision-making journey.**

### **Recommendation 8:**

- 9.3 A better way of supporting applicants to make informed choices would be to **bring together existing data and information sources in a more sophisticated and user-friendly online interface**. This could be done as part of the redesign of Unistats, and we would be happy to help test any proposed models.

### *Activity to enhance learning and teaching in individual subjects*

- 9.4 Russell Group universities are already dedicated to enhancing their provision for the benefit of their students and are delivering on this commitment in a wide range of ways including (but not limited to) the following:
- funding and undertaking new research into teaching and learning through internal teaching innovation funds and in-house teaching-focused journals

<sup>18</sup> <https://assets.publishing.service.gov.uk/media/53355970ed915d630e000017/OFT1529s.pdf>

<sup>19</sup> [https://www.heacademy.ac.uk/system/files/resources/student\\_choice.pdf](https://www.heacademy.ac.uk/system/files/resources/student_choice.pdf)

- recognising and rewarding contributions from academic staff through promotions criteria to support teaching staff to progress, annual teaching awards, and fellowship and CPD opportunities
- use of technology-enabled learning to maximise the value of face-to-face interactions between students and academics, and collection of data through learning analytics to improve learner outcomes
- investment in learning and teaching resources including libraries, museum collections, teaching spaces, labs, IT services, and subject-specific resources.

9.5 There are a number of mechanisms in place which will continue to drive improvements in quality including internal quality assurance mechanisms, the external examiner system, and extensive audit and reviews by a wide-range of Professional, Statutory, and Regulatory Bodies (PSRBs). In addition, the fees-based system and removal of student number controls mean universities are operating in an increasingly competitive market and are seeking to meet students' heightened expectations about the quality of the experience they can expect to receive. In this context, we do not believe a subject-level TEF assessment exercise is necessary on top of the provider-level TEF as a mechanism to encourage institutions to prioritise the enhancement of teaching and learning.

## 10. Further development of provider-level TEF

10.1 As outlined in sections 3, 4 and 5 above, **there are a number of reforms we recommend making to provider-level TEF** including:

- placing substantially more weight on absolute values alongside benchmark<sup>20</sup> and sector average scores
- removing the flawed 'grade inflation' metric and ensuring institutions subject to NSS boycotts or action which could skew NSS results are not penalised
- replacing the blunt medal rating system with a "profile approach" which better recognises the diversity of the sector and provides additional information for prospective applicants.

10.2 Whilst these reforms do not entirely mitigate our concerns, they should help to strengthen the usefulness of provider-level TEF as a tool for prospective students and for institutions. In further developing the provider-level TEF, it would also be useful to consider whether it would be appropriate to re-brand the exercise to reflect that it has changed. The current title of the exercise places an undue onus on teaching, given that the exercise does not involve any direct assessment of teaching.

### Recommendation 9:

10.3 Academic buy-in is vital to the integrity of any national framework which assesses the quality of higher education yet it is clear academics do not currently see the value of the TEF and research has shown they do not regard the existing methodology to be fit for purpose.<sup>21</sup> **The further development of TEF should be undertaken in collaboration with higher education institutions and academic staff in order to build buy-in for the exercise from the academic community.** We also recommend that the TEF is periodically reviewed to ensure it can evolve in line with developments across the higher education sector.

<sup>20</sup> Again, subject to caveats on benchmarking flaws that will need to be addressed

<sup>21</sup> <http://www.ucu.org.uk/media/10092/Impact-of-TEF-report-Feb-2019/pdf/ImpactofTEFreportFeb2019>

- 10.4 This could be undertaken through consultation and piloting in the same way that the RAE/REF was developed through a number of cycles over an extended period. A formal mechanism to ensure co-design is built into the further development of the TEF would be helpful; for example, the creation of a sector standing committee with responsibility for the development of TEF.
- 10.5 There are particular concerns about the validity of some of the TEF metrics, the problems they present when applied to the highly flexible four-year undergraduate degree structure in Scotland, and the added value they bring to the existing teaching quality framework applied in Scotland. Similarly, HEFCW in Wales requires all HEIs to submit an annual fee and access plan and TEF does not align with this process at all.

**Recommendation 10:**

- 10.6 While the development of the TEF has predominantly been focused on its application within English higher education, engagement with institutions and sector agencies in the devolved nations should be an integral part of any further work to develop TEF. This should reflect on the different regulatory systems in place across the devolved nations and enable consideration of whether and, if so, how the framework could add value to existing quality assurance regimes in the devolved nations. In particular, further consideration should be given to the way in which TEF is communicated domestically and internationally to ensure applicants recognise that participation is optional for providers in the devolved nations. The decision taken by providers in Scotland, Wales and Northern Ireland on whether or not to engage with the TEF should not be regarded as indicative of the quality of these institutions.

February 2019