

# Independent Review of the Teaching Excellence and Student Outcomes Framework (TEF)

## Call for Views

Launched on 18 January 2019

Respond by 1 March 2019

Do you support the aim of assessing the quality of teaching excellence and student outcomes across providers of higher education? Please explain why.

The Government has stated that the **purpose** of the TEF is to:

- better inform students' choices about what and where to study;
- raise esteem for teaching;
- recognise and reward excellent teaching; and
- better meet the needs of employers, business, industry and the professions.

We would support the aim of assessing the quality of teaching excellence and student outcomes across providers of higher education. However, the TEF in its current form does not measure teaching quality but rather uses proxy measures to assume this. The data becomes more problematic at a subject level when assessing outcomes across the sector due to the increased amounts of missing and/or suppressed data. Given the limitation of the data, it is therefore very important to retain the qualitative evidence and to place weight and importance on the peer review process. The aim of assessing teaching quality is a laudable one if it can be achieved.

1. These purposes fall into two main areas: providing information, and enhancing the provision of higher education.

a. Which of these is the most important (select one option only)?

- Providing information
- Enhancing provision
- Both are equally important
- Neither are important

b. Please outline below the reasons for your answers

The current aims of TEF to both provide information to applicants and drive enhancement do not always complement each other. If the TEF submissions (and panel judgements) are written with providing students with (additional) information in mind, there is a risk that some providers will use the submissions as a marketing exercise. This would result in providers being less open around areas of their provision which may need enhancing and how this enhancement might take place. Surveys of prospective students have consistently revealed that applicants are not aware that the TEF exercise is benchmarked and do not understand what the implications of this benchmarking are. The benchmarking is therefore not conducive to providing applicants with easily understandable information and indeed, is potentially misleading as lower TEF rated institutions can have better outcomes than higher rated TEF ones. This is not presented to prospective students in a way that means they are aware of this. There is a risk that the TEF could be regressive given the misleading nature of ratings for students. Students from more disadvantaged backgrounds are potentially more likely to be misled given the greater likelihood that they and their families will have less familiarity and context

regarding higher education. Prospective applicants already have access to a wide range of information on institutions they may wish to apply to. The DfE TEF2 evaluation report states that from their student surveys and interviews, there was no evidence that using the TEF contributed to higher levels of applicant satisfaction with the information available to help make decisions.

TEF should be used to enhance provision. It has already proved useful to the sector by focusing attention on gaps between different types of student through the split metrics. The benchmarking can be useful for institutions as it highlights differences between groups and distance from benchmarks which encourages enhancement even where an institution may have high overall values in absolute terms. The benchmarks can be beneficial for institutions but are misleading for students where absolute values matter. Given the information already available to students, contributing positively to enhancing provision and helping to raise esteem and rebalance between teaching and research is a more important aim of TEF.

## 2. Should there be any other purposes for TEF?

No, the TEF should not try to have multiple purposes as it will make it more unlikely that it will achieve its aims if there are multiple aims for multiple audiences.

## What is TEF?

This section provides some information about how TEF is currently designed and being delivered to help inform your response. This overview of TEF should be considered when responding to the remaining questions in this call for views. In particular, when considering Terms of Reference 1 and 2 in the next section.

The TEF is a national exercise, introduced by the government in England and open to all UK higher education providers. It assesses excellence in their teaching and how well they ensure excellent outcomes for their students in terms of graduate-level employment or further study. Providers are given one of three ratings – Gold, Silver and Bronze – or a provisional award if they do not yet have sufficient data to be fully assessed.

The diagram in Figure 1 provides an overview of the current TEF scheme (and elements proposed for subject TEF) and the process by which the ratings are determined. This shows that TEF draws on a mixture of quantitative and qualitative evidence. This includes measures drawing on existing quantitative data, all of which are gathered for other purposes, and a submission, which is written specifically to support the TEF process.

Some of the technical elements shown in the diagram are described briefly here:

- **Independent Panel** – TEF awards are judged by an independent panel of assessors, including students, academics (experts in teaching and learning) and experts in employment and widening participation.
- **Metrics** – These are measures deriving from national surveys and data returns, which are a key part of the evidence used in TEF assessments.
- **Split metrics** – Each metric is split out into sub categories to show how a provider performs with respect to different student groups (for example, age, ethnicity, educational disadvantage etc) and/or in different years.
- **Benchmarking** – This is used to allow meaningful interpretation of a provider's metrics by taking into account the mix of students and subjects taught at that provider. The benchmark is a weighted sector average where weightings are based on the characteristics of the students at the provider. The benchmarking factors currently accounted for in TEF are subject of study, entry qualifications, age, ethnicity, sex, disability, educational disadvantage, qualification level and the year of data used.
- **Very high and low absolute values** – Absolute performance is marked in the metrics based on the top and bottom 10 per cent of providers for that metric.

For more information about TEF, please see:

- The Office for Students' (OfS) [What is the TEF?](#) page for a brief overview of TEF
- OfS' [TEF information](#) pages for detailed information and technical documents
- The Department for Education's (DfE)'s [TEF page](#), including the [TEF specification](#).

Figure 1: The TEF process and evidence base

### The TEF focuses on three aspects of quality

Teaching Quality (TQ)	Learning Environment (LE)	Student Outcomes and Learning Gain (SO)
-----------------------	---------------------------	---

### There are 11 criteria mapped to these three aspects

<ul style="list-style-type: none"> <li>• Student engagement with learning (TQ1)</li> <li>• Valuing teaching (TQ2)</li> <li>• Rigour and stretch (TQ3)</li> <li>• Assessment and feedback (TQ4)</li> <li>• <i>Student partnership (TQ5)</i><sup>1</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Resources (LE1)</li> <li>• Scholarship, research and professional practice (LE2)</li> <li>• Personalised learning (LE3)</li> </ul>	<ul style="list-style-type: none"> <li>• Employability and transferable skills (SO1)</li> <li>• Employment and further study (SO2)</li> <li>• Positive outcomes for all (SO3)</li> </ul>
---	---	--

### A three-step assessment process is undertaken by an independent panel

#### Step 1: metrics based initial hypothesis

Quantitative evidence (a set of standard metrics and data) is considered by assessors to form an initial hypothesis of the rating of Gold, Silver or Bronze. This evidence includes:

Contextual data about the location, size and student cohort at the provider and/or subject

#### Core metrics (produced by the OfS)

<ul style="list-style-type: none"> <li>• Teaching on my course - National Student Survey (NSS)</li> <li>• Assessment and feedback (NSS)</li> <li>• <i>Student voice (NSS)</i><sup>1</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Academic support (NSS)</li> <li>• <i>Learning resources (NSS)</i><sup>1</sup></li> <li>• Continuation - Higher Education Statistics Agency / Individualised Learner Record data</li> </ul>	<ul style="list-style-type: none"> <li>• Employment or further study - Destination of leavers from higher education (DLHE)<sup>2</sup></li> <li>• Highly skilled employment or <i>higher</i><sup>1</sup> further study (DLHE)<sup>2</sup></li> <li>• Sustained employment - Longitudinal Education Outcomes (LEO)<sup>2</sup></li> <li>• Above median earnings or <i>higher</i><sup>1</sup> further study (LEO)<sup>2</sup></li> </ul>
--	---	--

Benchmarking, Very High and Very Low Absolute Values and Split metrics are used to highlight positive and negative performance against the core metrics.

#### Supplementary data (produced by the OfS)

- In the subject-level pilots: grade inflation data and data on differential degree attainment by student background are being tested (at provider-level only)
- For provider-level TEF: grade inflation data and the two LEO-based metrics are considered as supplementary evidence (in step 2 alongside the submission)

<sup>1</sup> New elements being tested in the subject-level pilot <sup>2</sup> The subject pilot is testing the two LEO metrics as Core metrics, and omits Employment or further study. Provider-level TEF uses only the two DLHE metrics as Core metrics.

#### Step 2: additional evidence to test initial hypothesis

**Submission** – a document written by the provider, with student involvement, that includes additional evidence (qualitative and/or quantitative) against the criteria, is considered.

#### Step 3: holistic judgement

Assessors review all evidence and make a best-fit judgement against the rating descriptors.

### The outcome is one of these ratings and a statement of findings



\* given to providers that meet national quality requirements, but do not yet have sufficient data to be fully assessed

# How well does TEF work?

## Terms of Reference

- 1: The process by which ratings are determined under the scheme and the sources of statistical information used in that process**
- 2: Whether that process, and those sources of statistical information, are fit for use for the purpose of determining ratings under the scheme**

An independent expert view of the statistical information used in TEF will be commissioned separately, but we would also like your views on the following questions.

As referred to in the TEF overview (previous section), TEF is being delivered at both provider-level and subject-level. We are interested in views on both. Please make clear which level you are referring to in your answers to each question.

3. Are the criteria used in TEF (see Figure 1 for a list of the criteria) appropriate?

If not, what criteria would be more appropriate?

The criteria used in the TEF are appropriate. The addition of a student engagement with learning criteria which has been added for the second year of subject pilots will hopefully emphasise that a provider alone is not wholly responsible for its student outcomes as a student also must engage and contribute to their own learning.

4. There is no direct measurement of teaching quality currently available. As a result, the TEF uses existing data as indirect measures of teaching quality. These measures are known as “proxies”.

- a. Are the metrics used in TEF the best proxies for measuring the TEF criteria (see Figure 1 for a list of the criteria and metrics)

Using the NSS as a measure for teaching quality is problematic, particularly because students completing the survey will likely only have experienced one higher education provider and therefore cannot compare their experiences. This makes the survey highly subjective. Institutions also do not have complete control over the student experience and therefore cannot necessarily improve their NSS scores easily for example, for half of the medicine course, the students' teaching and placements occur in the NHS. Currently, the NHS is under tremendous stress which can impact upon our students. This is outside the institutions ability to change and yet this affects NSS scores and therefore TEF scores. However, we do agree that the TEF should incorporate student feedback as part of the measures.

- b. If you answered no, what metrics would be more suitable proxies?

The addition of a 'completion rate' measure could be another suitable proxy. This would take account of and recognize students successfully completing a qualification during their time in higher education. Continuation is a measure of continuing to the second year of study, which is important. However, achieving a qualification is also an important aspect of higher or

## How well does TEF work?

further education.

5. The TEF metrics are benchmarked to account for factors such as the subject of study, prior attainment, ethnicity and educational disadvantage of the provider's student intake (see that 'What is TEF?' section for detail).
  - a. Should the metrics be benchmarked to allow for difference in a provider's student population?

If the aim of the TEF is to provide information to students on prospective universities then, as previously mentioned, the benchmarking can be very misleading and could lead students to think that a provider has far worse student outcomes than it does in absolute terms.

There is also a problem presented with provider's contributing to their own benchmark, this is particularly true at subject level. If a provider teaches a large majority of the students studying a subject across the sector then it is highly unlikely to be able to exceed the benchmark when its own scores contribute disproportionately to that benchmark.

If providing prospective students with information remains an aim of TEF then providing benchmarking among groups of providers with similar entry tariffs (which would therefore likely be institutions an individual student would choose between) could be a way to make this meaningful to a student. If providers were grouped, then meaningful comparisons between them could be made as their absolute outcomes would be similar allowing for benchmarking differences to have meaning without being misleading.

- b. Does TEF benchmark for the right factors?

Including overseas ethnicity in the benchmarking can result in a high degree of suppressed data. At a sector level, due to overseas ethnicity being a non-mandatory field in the HESA return which means the data is frequently not returned, the data on overseas ethnicity is not robust and should not be used. Data such as ethnicity, especially overseas ethnicity, where there is a high degree of 'unknown/ information refused' answers does not allow meaningful interpretations to be made and therefore should not be included in the benchmarking. As any benchmark where 50% of the data is unknown is suppressed, providers with large numbers of overseas students (where ethnicity is often categorized as unknown) can have large amounts of their TEF data suppressed. This effect is especially pronounced at subject level where there are much smaller student numbers being benchmarked. This is an unintended consequence of the benchmarking that could result in providers that *do* have the requisite NSS data not being able to receive a TEF award due to data suppression that has only occurred as a result of benchmarking and a provider should not be penalized and receive no award because of this.

Location could be considered as an additional benchmark factor given the impact that it can have on the metrics such as NSS scores and salaries.

6. The TEF process uses both quantitative evidence (for example, the core metrics) and qualitative evidence (for example, the written submission).
  - a. What are your views about the balance of quantitative and qualitative evidence

## How well does TEF work?

considered in arriving at ratings?

Given that the metrics behind the TEF are acknowledged to be proxies of teaching quality, it is imperative that the process gives equal emphasis to the qualitative evidence and to the independent peer review process which determines the ratings. The panels need to be able to use their own judgement in coming to a rating and the ranking criteria must therefore not be too prescriptive. The panel reports from the first year of subject pilots highlight that some of the panels did not feel empowered to make such decisions.

The qualitative evidence is very important. Institutions need the opportunity to demonstrate the ways in which their outcomes are linked to teaching quality given that the metrics being used are proxies. The submissions therefore give providers the ability to provide additional context and explain what their teaching methods and pedagogy is and why these work and link back to the metrics.

b. Are there any other aspects of the process that you wish to comment on?

No.



## Are the ratings right?

### Terms of Reference 3: The names of the ratings under the scheme and whether those names are appropriate

There are currently three ratings – Gold, Silver and Bronze – plus an additional Provisional award, given to participating providers that meet national quality requirements, but do not yet have sufficient data to be fully assessed.

#### 7. Are the purpose(s) of TEF met by:

##### a. awarding a single rating?

If the intention of TEF is to enhance provision for providers, then yes. Students' needs are not well served by a single rating. This view is supported by the applicant feedback in the DfE TEF2 evaluation where applicants stated that understanding why a certain award was given and more information on how different aspects of the award and assessment were made would be helpful.

##### b. with three levels of differentiation, plus a fourth rating for those unable to be assessed?

If one of the aims of TEF continues to be to provide information to students then consideration needs to be given to how a student interprets a rating of "unable to be assessed". There would need to be an explanation as to why a rating could not be provided. The DfE evaluation of TEF year 2 highlights that applicants are uncertain regarding the status of Provisional Awards which would support providing an explanation. The methodology of the TEF should aim to minimize that number of institutions that receive a provisional rating or are unable to be assessed given the reputational damage that such a rating could have. For example, the benchmarking should not lead to metrics being suppressed without a good reason (e.g. data protection of small numbers). As previously mentioned, including overseas ethnicity leads to unnecessary data suppression which could result in not receiving a rating. This should not happen and the OfS should mitigate where possible to ensure that as many institutions as possible can receive a rating.

##### c. ratings named Gold, Silver, Bronze and Provisional?

Please explain your answers.

Using Gold, Silver and Bronze are emotive ratings. A bronze rating is still intended to show teaching which is above national quality requirements (as is a provisional rating). However, this is often not how the awards are interpreted in practice. There is a concern that using this terminology may be damaging to the sector's reputation potentially especially with overseas audiences.

#### 8. If you answered no, what alternatives you would suggest:

##### a. For provider-level TEF?

A very short textual description of a few words could be used instead. Alternatively, a numerical rating could be used, especially if this was used where the highest number

## Are the ratings right?

represented gold e.g. gold is 3, silver is 2 and bronze is 1; this would make the ratings less emotive.

- b. For subject-level TEF?

Please see response to 8a.

- c. If your previous response(s) reflected on the impact of the TEF on the international reputation of institutions and/or the UK as a whole, we would welcome any evidence or information you can provide that might support your view or help inform the independent review.

Research undertaken by the Students Union Research Group highlighted that 38% students would reconsider applying to their course if it had been given a bronze rating and 14% would not have applied to a bronze course at all. The number who would still have applied to a bronze rated course is 36% if they had known it was bronze before applying. This implies that a rating below silver would have a negative reputational affect. As previously highlighted, this research also found no evidence that students understood that the TEF Medals are based upon performance against a benchmark rather than absolute performance: whenever students talked about medals, they consistently assumed that a 'Gold' rating at one institution is directly comparable to the 'Gold' rating at another institution. This lack of understanding of benchmarking is highlighted in the DfE TEF2 evaluation which notes that applicants commented on experiencing confusion around the polarised results between TEF and league tables where a highly ranked University would have a bronze or silver TEF award.

The DfE evaluation of TEF2 highlights that 'applicants displayed unprompted misconceptions about the TEF... for example [that] bronze awards indicate 'under achievement', rather than an achievement beyond satisfactory practice'.

## Has TEF changed anything?

**Terms of Reference 4: The impact of the scheme on the ability of higher education providers to which the scheme applies to carry out their functions (including in particular their functions relating to teaching and research)**

The review will consider the recently published [Evaluation of Provider-level TEF 2016-17 \(Year 2\)](#) as well as other available evidence, but we are also interested in your thoughts.

9. Has the introduction of TEF **positively** changed the educational experience of students (e.g. teaching and learning)?

If yes, how?

The TEF has focused attention on gaps and differences between certain groups not just on overall attainment which is positive. The use of benchmarking also means that providers are aware of their performance in relation to benchmark and this can drive enhancement and improvement despite high overall scores.

10. Has the introduction of TEF **negatively** changed the educational experience of students (e.g. teaching and learning)?

If yes, how?

There has been no negative affect of TEF on the educational enhancement of students. Subject level TEF requires considerable input from academic staff who have teaching responsibilities. The exercise therefore could take time that these staff would otherwise be using to devote to the teaching and learning experience. Whilst reflecting on practice and pedagogy is positive, TEF can still be a large piece of work for institutions over tight timeframes. There is also the potential for unintended consequences where providers could chase performance in metrics rather than focusing on the fundamentals of pedagogy, for example, by making teaching less stretching to improve satisfaction scores.

11. Has the introduction of TEF impacted **positively** on research and/or knowledge transfer?

If yes, how?

No.

12. Has the introduction of TEF impacted **negatively** on research and/or knowledge transfer?

If yes, how?

## Has TEF changed anything?

There have been no negative impacts thus far. However, the possibility that the TEF and REF cycles may occur concurrently, and that academic staff will be called on for their expertise for both exercises means that the TEF may negatively impact on research and knowledge transfer moving forwards. This is particularly true if similar staff are being called on to participate in the assessment panels of TEF and REF.

## Is TEF worth it?

### Terms of Reference 5: An assessment of whether the scheme is in the public interest

We are interested to assess whether the scheme provides outputs that are in the public interest and arrives at these outputs in a cost effective way that meets public interest principles.

The review intends to consider research about the costs and benefits of provider and subject level TEF, but we are also interested in your views of the range of benefits, and costs, of the scheme to individuals, institutions and society.

13. Does TEF help you as a student/student union/provider/employer/other? Please explain the reasons for your answer.

Yes. The process of the exercise is helpful as a provider especially with regards to performance relative to benchmark and the split metrics. However, being a gold rated institution has not necessarily been 'helpful' or changed applicant behavior.

14. Explaining your reasoning, what are the most significant **costs** of:

The most significant costs for both elements of the TEF are the staff costs involved in preparing the submissions for the TEF. The costs are considerably higher at subject level given the increased number of staff required for the submission. The subject level TEF requires significant input from academic staff time which makes this exercise far more costly and also increases the opportunity cost of the exercise. There is also a cost of academic time (and opportunity cost) on the assessment panels.

a. Provider-level TEF?

b. Subject-level TEF?

15. Explaining your reasoning, what are the most significant **benefits** of:

The most significant benefit of both elements of the TEF is the focus on teaching and rebalancing the emphasis between recognizing excellent teaching and recognizing excellent research. As described earlier in our response to Q10, the main benefits of TEF are the information providers receive on differences in student groups.

a. Provider-level TEF?

b. Subject-level TEF?

## Terms of Reference 6: Any other matters that the appointed person considers relevant

The review will also consider whether the overall TEF process delivers effectively for all across a diverse sector.

In the following question, we are particularly interested in views about:

- providers in Scotland, Wales and Northern Ireland;
- harder to reach applicant/student groups;
- part-time students and part time provision;
- small providers;
- specialist providers and specialist provision;
- private providers;
- further education colleges providing higher education.

16. Are there particular types of students, provision or providers that are **disadvantaged** by the current design of TEF, in a disproportionate

way? If so, what changes could be made to address

this?

Please see responses to questions 2b and 6 regarding how providers can be disadvantaged due to contributing to their own benchmark.

As previously mentioned, providers with large numbers of overseas students are also disadvantaged as their data is more likely to be suppressed due to overseas ethnicity not being a mandatory requirement of the HESA return.

In addition, the current level of knowledge needed to interpret and understand the TEF, in particular by prospective students, needs to be considered. It is reasonable to assume that those students, and their families, with less familiarity with higher education will be at a disadvantage in accessing and interpreting this information. Careful consideration will need to be given to how far it will meet its objective of providing useful information for effective student decision making.

17. Are there particular types of students, provision or providers that are **advantaged** by the current design of TEF, in a disproportionate way?

If so, what changes could be made to address this?

Please see the response to Q17.