Response ID ANON-SEUV-HDEQ-5

Submitted to **Teaching Excellence and Student Outcomes Framework: Subject-level**
Submitted on **2018-05-21 09:17:46**

Introduction

i What is your name?

Name:
Kevin Houston

ii What is your email address?

Email:
education@lms.ac.uk

iii I am a:

Representative organisation, business, or trade body

Please state:

iv If applicable, what is your organisation's name?

Organisation:
London Mathematical Society

v Would you like us to keep your responses confidential?

No

Reason for confidentiality:

Subject classification system

1 To define 'subjects' in subject-level TEF, do you:

Yes - agree

If you answered No, what other systems could be used and why?:

No

If you answered Yes, please explain why.:

Duration of award

2 Do you agree that we should have a longer duration and re-application period in subject-level TEF?

No - disagree

The focus of this question is on whether we should extend the duration. However, please provide as much detail as you can on your preferred length for the duration and/or re-application period.:
A subject level exercise that requires more than a single year to operate might in itself make that exercise not fit for purpose, as this would clearly indicate that the administrative burden of undertaking the exercise is considerable, and perhaps excessive.

Overview of subject-level TEF design

3 Should subject-level TEF retain the existing key elements of the provider-level framework (including the 10 TEF criteria, the same suite of metrics, benchmarking, submissions, an independent panel assessment process and the rating system)?

No - disagree

If you answered No, please explain why.:
The current provider-level system lacks transparency and did not receive the confidence of the sector. The resulting classifications seemed unfair and random to many. To carry this over into a subject level TEF is unwise. Complaints partly arose from the crudity of the output. Three possible outcomes are insufficient for providers to feel that they have been graded fairly. Furthermore, there is still the issue that some of the metrics do not measure what they are intended to measure. There is evidence that female lecturers receive lower scores than male

(https://www.timeshighereducation.com/news/growing-evidence-anti-female-bias-student-surveys) and that it can be many years after graduation that students appreciate the teaching methods used. For example, developing independence in students can easily be mistaken for failing to provide help.

Having different metrics at the institutional and subject levels will complicate the system and make difficult the comparisons required for each of the two proposed methods.

4 For the design of subject-level TEF, should the Government adopt:

A 'bottom up' approach (ie a form of Model B)

Please explain your answer. When answering this question, please consider the underlying principles that define Model A (a 'by exception' approach) versus model B (a 'bottom up' approach), and which principle you think we should adopt for subject-level TEF. While we are also interested in detailed comments on the specific design of each model, the final design will likely be a refined version of those presented in the consultation document. This question is therefore seeking views about which underlying approach you prefer. In your response, you may wish to consider the evaluation criteria set out in the specification for the first year of pilots (see below).:
Teaching within providers is generally done at the subject level with a greater variation of teaching methods and standard practice between subjects than between providers. Most students choose a subject first and a provider second so the information given to students should be as closely tailored to the subject as possible. This indicates that Method B is most suitable as then subjects guide the "overall" teaching quality indicator.

Furthermore, Method A clashes with the idea of allowing "distribution of subject ratings to vary naturally for each subject" as in section 10.3.

In both cases there is a problem of how to deal with borderline cases. This is particularly acute with Method B which is similar to the way in which degree classifications are awarded.

Model A: Generating exceptions

5 Under Model A, do you agree with the proposed approach for identifying subjects that will be assessed, which would constitute:

No - disagree

If you answered No, please explain why. You may wish to comment on variations or options that we have not mentioned:
Presumably providers will attempt to include as many exceptions as possible unless there is a downside to entering them. Experience from previous quality assessments (eg RAE) show that the system will be changed regularly as providers learn how to play the system. The area of exceptions is likely to be one such place where game-playing will occur.

The choice of method to trigger an exception is always going to be subjective. (No credible evidence is given for the planned method. It is entirely arbitrary.) Selecting Method A will lead to many departments within providers being unfairly gave a lower rating.

If a metrics driven system is used, then it makes more sense to let the figures speak for themselves

No - disagree

Please explain your answer. You may wish to comment on options for identifying the number of additional subjects or on any variations or options that we have not mentioned.:
See previous answer.

Model A: Relationship between provider and subject assessment

6 In Model A, should the subject ratings influence the provider rating?

Neither agree nor disagree

Please provide as much detail as you can on why and how this relationship should be brought about.:
That would seem to negate the idea of Method A.

Model B: Relationship between provider and subject assessment

7 In Model B, do you agree with the method for how the subject ratings inform the provider-level rating?

No - disagree

You may wish to comment on the method for calculating the subject-based initial hypothesis, as well as how this is used in the assessment process. We also welcome alternative approaches that do not use the subject-based initial hypothesis.:
Different subjects are likely to have different outcomes in terms of the number of Gold, Silver and Bronze outputs. A provider with a particular focus may benefit or be disadvantaged by such inconsistencies. Recall also that the Teaching Quality Assessments of 20 years ago were undermined by some subject area panels awarding full marks regardless of quality.

Metrics

8 Do you agree that grade inflation should only apply in the provider-level metrics?

Neither agree nor disagree

If you are able, please provide information about how grade boundaries are set within institutions to inform whether our rationale applies consistently across the sector. Comments on the potential impacts of applying grade inflation only at provider-level are also welcome.:
A problem with the grade inflation metric is that it lacks context and does not answer the question of why grades have increased. Grades could increase following improved teaching and learning or from lowering of standards. Nowhere is this adequately addressed.

Furthermore, a simplistic grade inflation metric does not take into account that grade distributions vary between subjects. Mathematics degrees result in more Firsts but also more Thirds. Using only a provider-level metric could distort the results.

9 What are your views on how we are approaching potential differences in the distribution of subject ratings?

You may wish to comment on our approach to very high and low absolute values, clustered metrics and regulation by Professional, Statutory and Regulatory Bodies (PSRBs).:
The proposed system of flagging does not seem to solve the problems.

The numbers generated through the TEF will be used by various bodies, particularly newspapers, to generate league tables. In the current system NSS scores are used but as many marks are clustered around the same value even small changes to an institution's subject score can dramatically change their position within a table leading to high year-to-year volatility. This is seen in Mathematics league tables and presumably happens for other subjects. It is not clear that the system addresses this.

10 To address the issue of non-reportable metrics:

Neither agree nor disagree

If you answered No, please explain why.:

rely on provider metrics alongside any reportable subject-level metrics?

Please explain your answer.:
In this situation giving assessors more information is likely to be helpful. This should be kept under review however.

Additional Evidence

11 Do you:

Yes - agree

If you answered No, please explain why.:

No

Please outline which subjects should have mandatory declaration and why.:
No opinion as the London Mathematical Society is concerned only with mathematics.

Interdisciplinarity

12 Do you agree with our approach to capturing interdisciplinary provision (in particular, joint and multi-subject combined courses)?

No - disagree

Please explain your answer. We want to ensure that providers are not discouraged from taking an interdisciplinary approach as an unintended consequence of subject-level TEF. We therefore welcome feedback on how the proposed approach will impact on providers and students.:
The alternative is to produce a complicated formula to combine information from different subjects. Determining this formula for different subjects would be very subjective and open to debate (and likely lead to unintended consequences involving subjects being dropped by providers).

However, the proposed method is unlikely to capture the necessary information on teaching since many providers provide separate modules for joint honours students. For example, a student taking Mathematics and Subject X may in their first year take mathematics modules that are mutually distinct from the modules taken by single honours Mathematics students. This may continue in subsequent levels. Hence, the information from the single honours teaching does not represent the quality of teaching in the joint honours subjects. Combining the information from the two may therefore be misleading.

Teaching Intensity

13 On balance, are you in favour of introducing a measure of teaching intensity in the TEF, and what might be the positive impacts or unintended consequences of implementing a measure of teaching intensity?

No - strongly disagree

Please explain your answer.:

We are very much against this proposal. It is an attempt to quantify something hard, perhaps impossible, to quantify.

The document concedes there is no correlation between hours taught and learning outcomes and teaching quality. Beyond a certain point more contact hours reduces student time for working by themselves and more important, particularly in mathematics, working with peers. The importance of the latter is often overlooked.

Nonetheless, teaching hours is essentially what forms the basis of most of the proposed methods. Attempting to improve the measure with the addition of class size and so on just makes it increasingly baroque and less related to what is intended to measure — teaching quality.

Adoption of such a measure will lead to distortions and unintended consequences as providers aim to increase their score by gaming the system.

14 What forms of contact and learning (e.g. lectures, seminars, work-based learning) should and should not be included in a measure of teaching intensity?

Question 14:

Regardless of what is included or not included, the system will be gamed to give an unrealistic picture of teaching quality.

For example, non-compulsory sessions are not so well attended by students. The provider could increase these without a commensurate increase in teachers needed and lead to distortion of the intensity figure. Excluding non-compulsory sessions penalises the providers who put them on to increase the quality of teaching provision to help lower-achieving students for example.

Another example: If the qualification of the teacher is not measured, then more post graduate students could be employed to run extra sessions. However, a PhD does not itself indicate excellence in teaching.

Similarly, what about peer assisted learning (PAL) sessions that are given by undergraduates from higher levels? These are incredibly valuable to learning. Include these in teaching intensity and providers may increase them to give a boost to the intensity score without a similar increase in quality.

15 What method(s)/option(s) do you think are best to measure teaching intensity? Please state if there are any options that you strongly oppose, and suggest any alternative options.

Question 15:

This is an attempt to measure with a single number a collection of interrelated concepts. All options are opposed. Each has its own downside which will warp teaching. Furthermore, such a metric discourages independence in students as providers as encouraged to provide more contact hours.

Introducing it is likely to stifle diversity as providers learn how to game it. Different providers have different philosophies. For example, contrast the approaches in the Open University, Oxbridge, Russell Group, and MillionPlus group universities, in terms of teaching mathematics. An intensity score will not reflect accurately what they do.

If one has to use crude metrics, then do not conflate them. Instead keep them separate. For example, stating contact hours and staff student ratios is already much better than inventing a new flawed concept.

Other comments

16 Do you have any comments on the design of subject-level TEF that are not captured in your response to the preceding questions in this consultation?

Question 16:

Using a graduate earning metric could lead to a perverse incentive that leads to long-term damage to mathematics. Providers will be incentivized to ensure their graduates earn as much as possible as this will be a key metric used in league tables. Since being a mathematics teacher is a relatively low earning career for a graduate mathematician it will not be in the interest of providers to promote this as a career. Clearly, this will of course lead to long-term damage to the teaching of mathematics.

The proposals will deter collaboration on teaching between institutions. Already some managers do not wish innovations to be shared with the "competition". The aim of the education system is to educate people. A system of competition will prevent good innovations in teaching from spreading and so perversely may lead to an overall loss to the UK's education system.