Teaching Excellence Framework: technical consultation for year 2

A response from the British Academy
July 2016

For further information, contact:
Jonathan Matthews
j.matthews@britac.ac.uk
+44 (0)20 7969 5214

INTRODUCTION

1. The Teaching Excellence Framework (TEF) should be considered in the context of the significant changes to the higher education and research landscape that are proposed in the Higher Education and Research Bill, further to the White Paper and Green Paper where the TEF was first outlined in detail by government. In its response to the Higher Education Green Paper[1] the Academy stated that the best teaching generally follows the best research. Assuming a false dichotomy between research and teaching is therefore unhelpful, and there is a risk that the sector will lose this oversight with the splitting of responsibilities for teaching and research in the Office for Students (OfS) and UK Research and Innovation (UKRI) respectively.

2. To turn to the detail of the proposed TEF mechanisms, the Academy is primarily concerned that the metrics that are intended to capture teaching excellence are fundamentally flawed. Crucially, no working definition of excellence in teaching is offered. The Academy has commissioned research to investigate the nature of teaching excellence in the Humanities and Social Sciences and would be pleased to discuss this commission with government further. Until there is a shared understanding of such a definition and an appropriate methodology for measuring it, there would be value in delaying the introduction of the TEF process.

DEFINITION (Question 1)

3. The TEF Technical Consultation states that its proposals have been developed to, among other aims, 'allow for diverse forms of excellence to be identified and recognised'. However, no definition is offered of what teaching excellence might comprise. The Academy has commissioned a team led by researchers at King's College London[2] to produce case studies of teaching excellence in HSS and would be pleased to discuss this commission further with government, as well as share the findings of the research.

4. There are several identifiable risks to the criteria proposed in Figure 4 under the headings 'Teaching Quality', 'Learning Environment', 'Student Outcomes and Learning Gain'.

5. Firstly, while weighted measures of class size are more appropriate than non-weighted measures, linking teaching quality too closely to contact time will not be a reliable measure of quality in all disciplines; it is better to focus instead on learning outcomes and experience. There is a great diversity of teaching approaches in higher education, involving more or less contact time. Increased contact time may not be the route to independent thinking and learning, which is a desirable outcome of any university teaching.

6. Secondly, the contractual status of staff involved in teaching is not a reliable indication of the quality of their teaching. For example, the Academy has produced

[1] http://www.britac.ac.uk/sites/default/files/British%20Academy%20Response%20to%20Higher%20Education%20Green%20Paper%20January%202015_0.pdf

[2] Teaching Excellence in HSS Commission http://blog.britac.ac.uk/teaching-excellence-in-the-humanities-and-social-sciences/

research[3] that demonstrates that postgraduate and postdoctoral teaching assistants for lab-work in quantitative skills works well, and is not simply a case of delivery at the lowest possible cost.

7. It should be stressed that, in the absence of robust and shared definitions of excellence, in practice providers will be driven by metrics. We go on to argue that no quantitative metric exists that can adequately capture teaching excellence across the great diversity of teaching and learning approaches and environments found in universities. It is also clear that the core metrics proposed based on the National Student Survey (NSS) questions are not fit for purpose, for both substantive and technical reasons.

METRICS (Questions 2-11)

8. The Technical Consultation argues for the diversity of forms that teaching excellence can take. The metrics it proposes, however, do little to capture that diversity. The core metrics proposed are based on NSS questions, which are not fit for purpose, for both substantive and technical reasons. This is both because student satisfaction and teaching quality are different things (so that the measures are of poor validity), and because the metrics as they stand do not differentiate among the vast majority of universities (so that there is very little variation).

9. There is a danger that the emphasis on provider-wide, student-satisfaction based metrics will discourage innovation in and provision of modules within degree programmes, for example on quantitative skills, if such modules, regardless of their pedagogical value, receive lower than average satisfaction ratings. Students revise their opinions about the relevance and quality of different components of their degree after a year or two in employment, often coming to see courses such as methods training, that they may have shown little interest or enthusiasm for at the time, as the most relevant for their careers in their degree.

10. Using multi-level analysis, Marsh and Cheng[4] show that only a very small percentage of the variation in student evaluations in the NSS are attributable either to universities or to subject areas within them. Most variation (about 9/10) is attributable to the individual student. That is, the variation in the way different students evaluate the same course swamps any average difference in the evaluation of different courses.

11. The ONS conducted a technical review of the NSS[5] that reaches a similar conclusion to Marsh and Cheng. They show that almost all the 95% confidence intervals for student ratings of courses within each individual university contain the mean for all universities: that is, most university performances are statistically indistinguishable.

---

[3] Measuring Up, http://www.britac.ac.uk/node/4188

[4] Herbert W Marsh and Jacqueline Cheng (2008) 'National Student Survey Of Teaching In UK Universities: Dimensionality, Multilevel Structure, And Differentiation At The Level Of University And Discipline: Preliminary Results', Dept. of Education, University of Oxford.

[5] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/523291/bis-16-269-teaching-excellence-framework-review-of-data-sources-interim-report.pdf

This is true when averaging the metric for all students at a university, regardless of course. If more precise comparisons are made, e.g. by gender or ethnicity or social class of students, or by subject of study, confidence intervals become much wider and virtually no sensible distinctions can be made.

12. Metrics such as those proposed are already being misused by third parties to produce 'league tables' that are fundamentally misleading. HEFCE's advice that it was a misuse and misinterpretation of NSS data to produce 'league tables' has not stopped this practice becoming widespread.

13. The Technical Consultation asks specifically about employment metrics. It is not at all clear that employability performance measures teaching quality as opposed to such factors as the skills with which students arrive at university, the reputation of the university, networking and other activities students undertake at university, or their subject choice. Employment destinations within a short period of graduation are a poor guide to later career progress. Data quality is an issue. The *Institute for Fiscal Studies* research by Britton et al (2016), even working with a 10% sample of administrative data from HMRC, found that data volume restricted what conclusions might be drawn. Useful and robust graduate employment outcome data would require timely and comprehensive data from the relevant government departments. Furthermore, employment or earnings do not exhaust the contribution of graduates to society, which may come about through voluntary work or many other forms of public service.

14. In addition to the core metrics, a provider submission will allow for additional evidence to be considered. The consultation document suggests that 'Assessors will be looking for evidence of how far a provider demonstrates teaching and learning excellence across its entire provision. The submission should therefore avoid focusing on successful but highly localised practices that affect a relatively small number of students studying on particular courses or in particular departments.' This may encourage providers to drive through 'one size fits all' teaching quality procedures that fail to recognise the particular nature and contribution of different disciplines or kinds of methodological training, for example in quantitative skills. It is also against the spirit of 'diversity of teaching excellence' that the TEF claims to endorse. One might expect improvements or step changes in teaching quality to be focused initially on particular units whose experiences others copy.

15. None of the foregoing precludes the extensive work that universities already do to monitor and manage the quality of teaching and learning that they provide. A 'portfolio' of such material would potentially be a better guide to teaching quality than any approach based on standardised metrics. Universities could be encouraged to publicise such material and alert potential students to it, through such mechanisms as the Key Information Set (KIS); many universities already do this.

16. Finally, due to the fundamental problems with the proposed metrics identified above, the TEF would be unlikely to identify low quality entrants to the market that are not focussed on providing teaching of genuine quality.