

Subject-Level TEF Consultation: Response from Goldsmiths, University of London

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

a) agree with using level 2 of the Common Aggregation Hierarchy as the classification system (CAH2, with 35 subjects), and if not, what other systems could be used and why?

We understand and agree with the principles behind the proposal to use CAH Level 2 subject classification. The subject breakdown aims to: be meaningful for students; group courses that are likely to be similar in teaching quality; have large enough cohorts to have robust data; and be manageable for providers and assessors.

However, we think that the further breakdown of some areas, in particular the 'Creative Arts and Design' subjects, is necessary in order to meet these principles. We would suggest that institutions should be able to make a case to make more than one submission against a Level 2 subject, as is permitted with UoAs in the REF, where there is clear justification in terms of distinctiveness of subjects, meaningfulness for students, and reportable metrics.

b) think that specific changes or tweaks need to be made to the definition of the 35 subjects in CAH2, or to the 7 subject groups used in Model B, and if so, please explain why?

We propose that the breakdown of 'Creative Arts and Design' should be beyond Level 2. This subject area is wide ranging, incorporating art, design, music, drama, dance, cinematics and photography - a diverse collection of disciplines requiring distinct delivery and specialist facilities.

For Goldsmiths, this Level 2 subject grouping covers four very different departments – music, theatre and performance, art, and design - teaching entirely distinct subjects. A single TEF rating for these departments would be of no more meaning to students than an institutional rating, and setting out the different approaches coherently within a single TEF submission would be challenging.

The Subject-Level TEF consultation documents show that only 9% of students fall into the Creative Arts and Design subject area, which may pose problems if numbers are broken down too far in terms of reportable data. However, given the specialist nature of the delivery of these disciplines, and the objective of the exercise being to better inform students, we feel there is a strong case for enabling the breakdown of this subject area to something closer to (though not necessarily quite as far as) Level 3.

Alternatively, the challenges outlined above could be addressed if we had the option to make more than one submission against the subject.

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

Goldsmiths does not agree with the proposal to extend the duration of award to five or six years, or with the proposal that opportunities to submit for assessment should be bi-annual (Option 1).

Any system that is eventually put in place should be designed to encourage and incentivise change. It is important that the opportunity to re-apply is available once there is institutional confidence that there is sufficient evidence to warrant a different rating. Change takes time and things are unlikely to shift vastly from one year to the next, but – and this is the critical point - institutions' rate of change will vary. To not have an annual assessment process may unfairly impact those institutions who feel able to demonstrate change outside the timeframe that the cycle allows.

There is concern voiced in the consultation documents that institutions that don't like/agree with their award will simply keep submitting for new assessments every year. We would argue that this is unlikely to be the case as it would be a heavy demand on resources for the institution and repeated failure to show improvement will carry reputational risks. It might be feasible to set parameters for re-submission, such as a minimum of two years between submissions, but we believe it is important that an annual process is available so that it is positive change rather than the cycle that drives submissions.

It is also worth noting that change can mean a fall in performance as well as an improvement. The concern that annual assessment exercises may tempt institutions to focus on 'gaming' the exercise rather than achieving genuine improvement could be countered with an argument that Gold awards for a duration of 5 or 6 years may lead to complacency and dips in performance rather than incentivising sustained improvement.

Awards that endure for 5 or 6 years would also be based on increasingly out-of-date information and data. The metrics are already retrospective; in its 6th year, for example, a TEF rating will have been assessed on metrics that relate to 9 or 10 years previously.

The only argument for lengthening the duration of awards and limiting the number of assessment rounds is to reduce the burden of undertaking TEF submissions. But, if the TEF exercise is meaningful and important and designed to drive change, then DfE / OfS shouldn't be afraid of requiring institutions to fully participate as part of their standard practices. Institutions will find ways to manage the cycle of preparing and undertaking submissions, embedding it into the usual cycle of monitoring and improving, as they do other exercises that are deemed important and meaningful.

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)?

Subject level TEF should remain in step with provider-level TEF, for clarity. There are many elements of the design of provider-level TEF which are problematic, including the crude division into three 'medal categories', the use of benchmarking which fails to take into account locality, and the heavy reliance on NSS scores as a proxy for 'quality'. We hope that all of these areas will be examined in the statutory review.

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

- A 'by exception' approach (i.e. a form of Model A), or
- A 'bottom up' approach (i.e. a form of Model B), or
- An alternative approach (please specify)?

Both models are flawed. Of the two, Goldsmiths believes that Model B best meets the main intention of Subject-Level TEF, i.e. providing greater information to students, as it allows for greater variation between outcomes for subjects. It also provides for a clearer, more transparent way of understanding the process by which a subject achieves a rating.

Model A, meanwhile, will see a greater proportion of subject ratings overtly influenced by institution-level metrics and therefore provide only limited additional information for students.

The primary argument in favour of Model A is that it will reduce the burden of submission and assessment. Whilst proportionality and value for money are of course important considerations, designing an exercise that focusses on minimising the burden for institutions and assessors rather than allowing for true differentiations to be visible to students can only undermine it.

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

a) the initial hypothesis rule for generating exceptions from the metrics?

Goldsmiths queries the fairness of the way in which Model A determines whether a subject is 'an exception' and therefore warrants a distinct submission. For a subject to be deemed an exception, its metrics must "differ sufficiently" from the institution-level metrics. However, this does not include the case where an institution has a negative or positive flag but the subject is performing within benchmark and therefore 'neutral'. In an institution with initial hypothesis Bronze metrics, any subject areas that were neutral would not be deemed sufficiently exceptional for a distinct submission, and would therefore be awarded the institutional Bronze rating automatically (assuming no shift in rating post the initial hypothesis).

On the other hand, if those subjects were in an institution with a Gold initial-rating, they would be automatically rated Gold.

In either case it is not clear whether the rating is deserved or the process by which that rating was awarded sufficiently transparent.

The only reason we can see for not recognising going from positively or negatively flagged to neutral as an exception is to limit the number of areas an institution would need to make submissions for, but this penalises the subjects performing within benchmark in Bronze institutions and hides differentiation in Gold institutions.

b) allowing providers to select a small number of additional subjects?

Goldsmiths would argue that institutions should be able to submit for any number of additional subjects if they so choose. Institutions are best-placed to determine if there is a case to be made that differentiates a subject area, and they should be trusted and empowered to make any decisions after balancing the burden of submission against the likelihood of achieving a differential outcome.

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

The proposed 'feedback loop' that allows the institution rating to be reviewed once subject assessments are complete would mean that institution-level ratings are influenced by changes in subject assessments, which would seem fair.

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

The method of weighting subject assessments according to the number of students in those subjects to inform the provider-level rating is clear and fair. The proposed rules around how Gold, Silver, and Bronze ratings will be applied at provider-level raise questions about the value of the other elements of the provider-level assessment, however.

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

Goldsmiths questions the inclusion of retrospective grade inflation metrics at either institution or subject level. There are other means by which standards and equity of assessment are monitored (e. g. QAA, External Examiners). The TEF exercise is designed to recognise excellence and drive improvement in teaching. That the shadow of suspicion will fall on institutions if this results in greater student success is difficult to rationalise.

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

The proposal to allow natural distribution of ratings within subjects is fair and transparent, and Goldsmiths supports this.

However, we are concerned about the proposal to use sector-wide thresholds to identify Very High or Very Low absolute scores in subject-level data. Whilst all other aspects of the subject-level assessment are based on subject-level metrics and submission, the use of non-subject-specific absolute values is statistically questionable, and potentially highly misleading for students. In a scenario where a subject is performing within or above benchmark, but is flagged a Very Low absolute because metrics for that subject across the sector are poor it is a) difficult to know how to address that in a submission from an institutional perspective and b) confusing and misleading for students and others to appreciate without lengthy explanation that a range of different things are being signalled.

High and Low Values are unlikely to drive students' subject choices - it is well-known that students decide first what subjects they want to study and then apply to different institutions based on degree content and a range of other factors. Given the overall purpose of informing student choice, it is more informative and transparent that all aspects of the metrics focus on understanding the relative performance of different institutions within the subject area concerned.

We understand the frustration around high benchmarks for some institutions and subject areas, however, given that the metrics are only initial indicators, it is relatively simple in these cases to draw attention to these factors within provider submissions.

10 To address the issue of non-reportable metrics:

a) do you agree with the proposed approach?

The proposed approach seems to be a reasonable and fair compromise between ensuring robust metrics and enabling institutions to apply for a TEF rating.

b) when assessment occurs, do you prefer that assessors:

- rely on group metrics alongside any reportable subject-level metrics?
- rely on provider metrics alongside any reportable subject-level metrics?
- follow an alternative approach (please specify)?

As a principle, we would argue that it is best for the metrics to be as close to the subject concerned as possible, and therefore would see group metrics as being more appropriate than provider metrics. Having said that, we are aware of instances at Goldsmiths where group level metrics would be as unrelated as provider level, given the distinctiveness of disciplines within some subject areas (e.g. Creative Arts, as noted above). This highlights a potential weakness of Subject-Level assessments, but we do not think that a more robust approach within the current framework is available at this stage.

11 Do you:

a) agree that QAA Subject Benchmark Statements and PSRB accreditation or recognition should remain as a voluntary declaration, and if not, why?

b) think that there are any subjects where mandatory declaration should apply?

The difficulties around introducing mandatory declarations of QAA Subject Benchmark Statements or PSRB accreditation are clearly set-out in the consultation documents. This does, however, give rise to potential tensions between different assessments of standards not aligning, and clarity about the differences will need to be very clear to students.

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

We understand and value the approach to enabling providers to choose to submit programmes that cross more than two subject areas into one of the three 'general' subject areas. Our only concern around this is with regard to students understanding these subject areas and which programmes fall within them. We foresee the need for

institutions to provide specific explanation to students in those cases where they opt to use the broader subject-areas.

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?

We understand the drive to identify a measure in this area, but we feel that there is a long way to go to develop something that is meaningful and robust.

Where there are no fixed definitions and means of recording and reporting of data between institutions then the robustness of that data must be challenged. We also query how meaningful a 'GTQ' score will be to students – a number would seem a far remove from the actual experience of teaching and learning.

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?

Goldsmiths would value the opportunity to learn more about the outcomes of the Teaching Intensity pilots in order to better understand and explore all the issues before forming an opinion.

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.

Goldsmiths would value the opportunity to learn more about the outcomes of the Teaching Intensity pilots in order to better understand and explore all the issues before forming an opinion.

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

There is a clear attempt in some of the approaches proposed to minimise the burden of undertaking TEF, and we fear that for those institutions who performed well in the initial provider-level exercise this may be a driving consideration in their responses to this consultation. It would be an interesting exercise, for example, to consider whether 'Gold' institutions are more or less likely to favour Model A than those with 'Silver' or 'Bronze' awards. We trust that an awareness of this self-interest will be taken into account when considering the outcome of the consultation.

More broadly, Goldsmiths sees the TEF as currently designed as a regrettable and clumsy mechanism far removed from the classroom and the intellectual endeavour that happens there. There is a huge range of data publicly available about the relative 'performance' of institutions and it seems unlikely that the addition of another form of rating – even one with Government authority – will provide students with any further information or clarity. Goldsmiths will continue to object to the TEF (even a

renamed one) on the principle that it fails to provide meaningful insight on the teaching and learning experience and to measure what it purports to.