



Universities Scotland response to UK Government consultation on subject-level TEF

Universities Scotland welcomes this opportunity to respond to the UK Department for Education (DfE) consultation on the design of the Teaching Excellence and Student Outcomes Framework (TEF) at subject level. This response outlines concerns in the Scottish sector about present proposals, with respect to important Scotland-specific issues as well as what we see as other fundamental flaws. In light of these reservations, it seems unlikely that there will be significant Scottish participation in subject-level TEF as currently conceived; however, the Scottish sector remains committed to working with colleagues across the UK to explore options for a fit-for-purpose subject-level approach.

Scotland's higher education institutions are committed to excellence in learning and teaching. The sector's rigorous enhancement-led approach to quality, rooted in student partnership, is well-established and widely-supported, and Scotland performs strongly in student satisfaction and graduate outcomes. Scottish institutions are also a core part of the UK higher education sector, with a stake in its international reputation for excellence, and we attract students from across the UK and beyond. The introduction of the TEF therefore has UK-wide implications, and it is vital that its design is informed by devolved nation perspectives.

To this end, the Scottish sector has engaged actively and constructively throughout the development of the TEF, and we welcome this opportunity to continue to contribute to this process. From the outset, the Scottish sector has had doubts about the prospect of implementing the TEF at subject level, and on the basis of the consultation proposals, we remain unconvinced that a credible subject-level approach which is workable in practice has yet been identified.

In our view, designing a viable model of TEF at subject level faces three key challenges:

- achieving meaningfulness for students and prospective students;
- securing the credibility of metrics at subject level; and
- ensuring proportionality of the burden on institutions.

For subject-level TEF to be valuable, it is vital that its outcomes are meaningful for students. Universities Scotland is concerned that several aspects of the proposals stand to undermine the usefulness of the subject-level exercise for students, including the broad



definition of subjects and the convoluted nature of the assessment process, and it is not clear to us whether there are workable solutions to this core problem.

We are also concerned that any subject-level model of TEF will involve assessment based on small numbers of students. This will lead to significant data problems and reduce the reliability of outcomes by limiting the availability of robust, reportable metrics, and reporting fluctuations that are not indicative of genuine, meaningful trends.

The proposed subject-level TEF models will be costly and burdensome for institutions and government, particularly with a granular approach to defining subjects. The value added for institutions and government is unclear, however. Other markers of quality exist at subject level already, and in Scotland the sector has a well-established process for subject review in place via its Quality Enhancement Framework.

The sector also has significant concerns about Scotland-specific issues. Scottish institutions offer four-year undergraduate degrees; at many institutions and in many subjects, this structure offers valuable flexibilities and provides a wide variety of possible degree pathways for students. At some institutions, it is common to study multiple subjects, and also to switch subjects, meaning it can be impossible to track a typical trajectory for students graduating with a specific degree, or define consistently which department a student “belongs to” throughout their degree. In addition, the range of entry and exit points available for degree study at some institutions via college articulation routes and ordinary degrees means that many students are never captured in NSS data, and this issue is particularly acute at subject-level.

Universities Scotland has previously supplied DfE with more detailed evidence on these distinctive features of Scottish higher education. We welcome the acknowledgement of this in the technical consultation document (p. 64), and the attempt to respond to data challenges arising from flexible and multi-subject provision. We are nevertheless not convinced that the consultation proposals fully address the Scotland-specific challenges we have outlined.

Our more detailed views on these issues, and others, are set out in response to the specific questions posed in the consultation document, below. Whether or not subject-level TEF is adopted, Universities Scotland looks forward to continued dialogue with DfE and colleagues across the UK as the framework continues to evolve, and also to offering input into the upcoming Independent Review of the TEF from a Scottish sector perspective.

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?

Yes – strongly agree



- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland is not convinced that it is possible for any classification system to strike a balance in the granularity of subject definition that provides meaningful value for students but does not simultaneously generate serious data problems and a significant burden for institutions and government. This challenge is not limited to the TEF, but creates particular issues for the TEF due to the varying purposes being served.

We believe that the proposed 35 subjects is insufficiently granular to ensure usefulness to students; it is likely to conflate a diverse variety of courses within institutions, and enable only limited comparability across institutions. It is unclear to us that the approach will add significant value over and above provider-level TEF.

At the same time however, 35 subjects is too granular ensure robust data. The technical consultation document itself highlights (p. 49) that by adopting CAH2, 126 HEIs will have non-reportable metrics in at least one subject. We believe this poses a fundamental challenge to the credibility of the exercise and its outcomes (and we are unconvinced by the proposals for managing non-reportable metrics, as noted in response to Question 10 below).

In addition to generating non-reportable metrics, the small numbers involved will reduce the reliability of outcomes and report fluctuations in data year-on-year that are not indicative of genuine, meaningful trends.

On these questions of the statistical validity of the subject-level TEF proposals, we find the commentary submitted by Professor John MacInnes (as part of the British Academy's input to the consultation) compelling.¹

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?

- Yes
- No

No comment (see our response to Question 1a above).

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

- Yes – strongly agree

¹ https://www.britac.ac.uk/sites/default/files/CommentarySubjectLevelTEFConsultation_JohnMacInnes.pdf



- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland believes that to alleviate the burden of subject-level TEF on institutions, a longer duration would be required. We would also caution against too long a minimum re-application period, however, which we believe would block institutions seeking an up-to-date assessment without a compelling rationale.

We would suggest an approach that falls in between the two options offered in the consultation document, with a six-year duration and a three-year minimum re-application period. (If the assessment process becomes two-yearly, as the consultation document notes it may, we suggest a six-year duration and two-year re-application period.)

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)?

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland has concerns that, as the TEF criteria included in the provider-level framework are not discipline-specific, these will not necessarily be equally applicable in all subject areas, or provide the most appropriate yardstick to support judgements about the quality of provision in a specific subject area.

We also continue to have reservations about several aspects of the metrics and benchmarking used in the provider-level framework, particularly the use of POLAR in benchmarking, the non-continuation metric, and the use of LEO data.

POLAR

As Universities Scotland has highlighted previously, POLAR is widely acknowledged to be less useful in Scotland and is not used by HESA, the SFC or the Scottish Government.

POLAR works by dividing the UK into areas according to HE participation. As Scotland has a higher rate of participation in HE than England, it has significantly fewer low



participation areas, meaning that it is significantly less likely that Scottish-domiciled entrants to Scottish institutions will come from such areas. As a result, SIMD is used instead of POLAR for TEF metric splits for Scottish institutions.

In Year Two, however, POLAR was used in the benchmarking process for the highly-skilled employment metric for all UK institutions, including those in Scotland, despite its unsuitability for this purpose.

HEFCE's own analysis in 2016 confirmed that including POLAR in benchmarking of the highly-skilled employment metric had a negative impact overall on the performance of Scottish institutions against benchmark, and identified a negative impact on the flags of at least two Scottish institutions. Moreover, the rationale for including POLAR in the benchmarking of this metric is based on DfE analysis that explicitly only examined a sample population of graduates from English institutions, so it remains unclear whether POLAR is, in fact, correlated with highly skilled employment outcomes for graduates from Scottish institutions at all.

Universities Scotland is disappointed that from Year Three, POLAR is intended to be used to benchmark three metrics for all UK institutions: highly-skilled employment, earnings above the median, and non-continuation. This approach does not support a level playing field for Scottish institutions, and we strongly encourage DfE to reconsider.

Non-continuation

In addition to the non-continuation metric now being benchmarked by POLAR, Universities Scotland believes that how retention data is calculated and subsequently captured in the TEF puts Scottish institutions at a disadvantage in at least two other ways. This is not a new issue, but becomes critical in the TEF due to the calculation of benchmarks using the UK-wide dataset, and the subsequent generation of flags.

Firstly, the TEF metric (and underlying HESA data) on retention does not capture transfers to colleges in Scotland. For a number of years, the percentage of students recorded as transferring from Scottish institutions to other providers after the first year of study has been consistently lower than in the rest of the UK. A key reason for this difference is the transfer of students from Scottish institutions to HE-level enrolment in Scottish FE colleges, which are not captured in HESA's non-continuation data used to generate TEF metrics. Initial work by the SFC in 2016 indicated that if students transferring to an HE enrolment in a college in Scotland were included in the 'continuing' figures, the 'no longer in HE' rate for the sector would fall on average by 0.9 percentage points (with a range by institution of -3.2% to 0%). This is not an exact model of the impact on the TEF non-continuation metric, but it provides a reasonable indication. The exclusion of college transfers in Scotland from the calculation of the TEF



metric is likely to lead to Scottish institutions on average having a higher non-continuation rate.

Secondly, the TEF metric is based wholly on data about non-continuation after the first year of study, and excludes other elements of HESA's UK Performance Indicators on retention, including resumption of study after a year out. On this measure, the Scottish sector has consistently outperformed the UK sector average for seven years in a row (i.e. as far back as data is available), suggesting that Scottish institutions are doing a consistently better job of supporting discontinuing students back on course. For example, in 2015/16, the proportion of students resuming study at the same institution after a year out was 14.8% for Scottish institutions, and 9.7% for English institutions. This success is not reflected in the single TEF metric on retention, however.

LEO

There are significant reservations in the Scottish sector about using the experimental LEO dataset within the TEF. As noted above, the relevant metric is benchmarked by POLAR, which is problematic for institutions in Scotland. Universities Scotland also has wider concerns about its shortcomings:

- **It implies success is a high salary.** The highest paid job is not necessarily the “best”. In recognition of this, the relevant metric measures only the proportion of graduates earning over the median salary, rather than salaries per se. In our view, this only partially mitigates the fundamental problem, however.
- **It ignores the role of learner choice.** There are countless reasons a graduate may consciously choose a career path that is less lucrative, but is nevertheless valuable to our society and economy, and meaningful and rewarding to that individual.
- **It ignores regional differences in salaries.** Average pay varies significantly across the different regions of the UK, notably being highest in London and south-east England; this is not taken into account in the data, which will disadvantage institutions in Scotland and other regions.
- **It ignores self-employed graduates.** LEO does not capture the incomes of graduates who are working on a self-employed basis. This is problematic for a sector developing enterprising graduates, and may have a particular impact on institutions with significant numbers of students choosing to pursue creative and entrepreneurial career paths.
- **It ignores differences in working patterns.** The dataset does not record employment in FTE terms, and this will have an impact on institutions with significant numbers of graduates who choose to work on a part-time basis. Also, we know disabled graduates are more likely to be in part-time work and less likely to be in full-time work than non-disabled graduates.
- **It ignores EU and international students.** As the dataset covers UK-domiciled students only, LEO overlooks Scotland's significant international student body



altogether, as well as any home students who choose to work abroad after graduation.

- **It is historical.** The dataset, by its very nature, looks backwards, involving a very significant time lag, and may not prove to be reliably predictive for recent graduates, let alone current or prospective students. For example, the cohort covered in the last LEO data release graduated as the economic downturn began.

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)?**

Universities Scotland has reservations about both approaches, and is unable to endorse either one. As already noted, however, the Scottish sector remains committed to working with colleagues across the UK to explore options for a fit-for-purpose subject-level approach.

Model A

We have concerns about the convoluted nature of the assessment process under Model A. The inconsistencies in the process – i.e. that the same subject is assessed in more detail at one institution than another; and that at the same institution, some subjects are assessed in more detail than others – are not reflected in the ratings awarded, and we believe this will undermine the clarity and meaningfulness of the outcomes and may mislead students.

The burden is likely to be high in many cases – institutions will be required to produce 15 pages of submission at provider level, plus 5 additional pages per exception subject. The 'by exception' approach will mean that the burden is unpredictable, however, which will present challenges for institutions in planning ahead.

The complex process for generating exceptions will exacerbate this issue, and may also mean that institutions are required to prepare subject-level submissions without a clear understanding of why the subject is identified as an exception whilst others are not. As noted in response to Question 5a below, we also have concerns about a reliance on an 'initial hypothesis rule' that is based wholly on metrics.

As noted in response to Question 1b above, 35 subjects will often be too broad to be genuinely useful to students or to fit departmental structures at many institutions, but will also be too granular to ensure robust data.



Model B

The seven broad subject groups used for submissions under Model B are problematic. The proposed groups combine diverse subjects in ways that will present significant challenges for many institutions in relation to fit with departmental structures and producing single submissions. It is unclear on what basis DfE believes that all subjects in each broad group “are likely to have similar teaching practices, teaching quality, and student outcomes” (p. 9), and we are not convinced that the maximum length of submission for each broad group will enable institutions to evidence excellence across a diverse range of subjects. At the same time, the burden will be consistently high for most institutions – up to 73 pages in total. We also have concerns that the proposals on varying the length of submissions by the number of subjects an institution offers may favour those with fewer subjects.

Like Model A, Model B also uses CAH2 for the purposes of generating metrics and awarding subject-level ratings.

Again, as noted in response to Question 1b above, 35 subjects will often be too broad to be genuinely useful to students or to fit departmental structures at many institutions, but will also be too granular to ensure robust data.

It is important that the process for using the subject-level ratings to generate a single provider-level rating is made clear and transparent to institutions and students, and does not disadvantage any institutions due to subject mix.

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?

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland has concerns about the reliance on an ‘initial hypothesis rule’ based wholly on metrics, particularly in light of the fundamental flaws in the robustness of the metrics under present proposals (as noted in our responses to Questions 1a, 3, 10 and 12).

In principle, we may be less concerned about this approach if subject-level metrics were underpinned by a more robust methodology. It may be nevertheless be preferable for the final provider-level rating to be generated first and for exceptions to be identified subsequently, to enable the identification of exception subjects to be informed by the



submission, and to enable institutions to produce subject-level submissions knowing the initial hypotheses.

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

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

As noted, Universities Scotland believes Model A is fundamentally flawed under present proposals. In principle, however, if a more workable formulation were to be found, Universities Scotland would be supportive of institutions being able to select additional exceptions, as we believe this could offer valuable opportunities for institutions to showcase areas of strength that are not identified by the metrics alone.

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

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

In the absence of a stronger evidence base informed by the outcomes of the pilot process, Universities Scotland does not believe we are in a position to reach a judgment on this issue.

In principle, we can see potential value in a feedback loop between subject-level and provider-level ratings, but we are not sure it would be possible to implement without further complicating an already convoluted assessment process under Model A.

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

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

In the absence of a stronger evidence base informed by the outcomes of the pilot process, Universities Scotland does not believe we are in a position to reach a judgment on this issue.



We believe it is important, however, that the process for using the subject-level ratings to generate a single provider-level rating is made clear and transparent to institutions and students, and does not disadvantage any institutions due to subject mix.

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

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland does not believe that a grade inflation measure should be introduced in the TEF at all, at provider-level or subject-level.

The Scottish sector has strong concerns about the potential impact of such a measure and the signal it would send. It seems highly likely that it would generate perverse incentives for institutions, and the implication – that an improvement in outcomes is assumed to be a result of falling standards – runs directly counter to the sector’s strong and long-standing commitment to an enhancement-led approach to quality.

In addition, implementation of this seems premature, in light of the fact that joint UK-wide work by UUK, GuildHE and QAA to examine the evidence on trends in grade improvement is ongoing.

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

We have some concern that differential distributions could lead to perverse incentives for institutions to adjust their subject mix to optimise their TEF ratings.

10. To address the issue of non-reportable metrics:

a) do you agree with the proposed approach?

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland has strong concerns about the prevalence of non-reportable metrics at subject-level due to the small numbers of students whom data is available at a granular level. We note that the consultation document itself highlights (p. 49) that by adopting CAH2, 126 HEIs will have non-reportable metrics in at least one subject.



This is a question of statistical validity of conclusions, and it is correct to suppress the data where it is insufficient to draw robust conclusions. However, this will lead to significant gaps in available information, and it is far from clear to us that the proposed approach to managing these non-reportable metrics will generate robust, comparable results. We believe that allowing a subject with only two reportable metrics to be assessed, whether by substituting data from the broad subject group or the provider overall, stands to undermine completely the credibility and comprehensibility of subject-level TEF and its outcomes.

However, even if a workable approach to managing non-reportable metrics could be developed, wider fundamental problems with the robustness of the metrics under the present proposals (as highlighted in our responses to Questions 1a, 3 and 12) still stand.

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)?**

No comment (see our response to Question 10a above).

11. Do you:

a) agree that QAA Subject Benchmark Statements and PSRB accreditation or recognition should remain as a voluntary declaration?

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland believes that QAA Subject Benchmark Statements and PSRB accreditation provide valuable additional evidence beyond the metrics, and that significant weight ought to be attached to this evidence when institutions choose to provide it. We do not believe it is necessary to be prescriptive about this.

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

- Yes
- No

We do not believe it is necessary to be prescriptive about the additional evidence provided by institutions.



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

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland has significant concerns about the proposals to use ‘general’ subjects in CAH2 to capture multi-subject provision, which is common in many Scottish institutions, particularly during the first two years of the four-year degree. The flexibility in subject choice is reflected in a large number of programmes at many institutions. University of Edinburgh, for example, has almost 500 distinct programmes, of which 62% have at least two different JACS codes, whilst the University of St Andrews has over 900 distinct programmes. We cannot see how categorising a high proportion of these programmes under ‘general’ subjects could possibly provide any meaningful value to students. Moreover, we are concerned about potential reputational risk to the Scottish sector if much of its provision were simply to be labelled as ‘general’ to prospective students.

Another key Scotland-specific challenge is the high rate of students at many institutions who switch programmes during their studies and graduate with a different degree from that commenced in first year. For example, at the University of Glasgow and the University of St Andrews, around half of students do so. This results in a more indirect link between entry and exit, reducing the relevance of subject-specific outcomes data for a significant proportion of students. We do not see how the consultation proposals address this matter at all.

In addition, at some institutions there are high levels of students entering degree study in Year Two or Three via articulation routes from colleges, as well as high rates students exiting with an ordinary degree after Year Three to pursue employment opportunities. These are valid routes into and out of degree study in Scotland. Many students on such pathways are never captured in the NSS data, however, and at subject level this issue is particularly acute.

More detailed evidence on these distinctive features of the Scottish higher education context has been supplied to DfE separately, and the Scottish sector remains open to ongoing dialogue to continue to explore these matters.



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?

- Yes – strongly agree
- Yes – agree
- Neither agree nor disagree
- No – disagree
- No – strongly disagree

Universities Scotland is not in favour of introducing a measure of teaching intensity in the TEF.

The method currently being tested via the subject-level pilots is crude and yet highly costly for institutions, and we do not believe that any credible alternative exists. The options presented in the consultation document are not realistically workable, nor are they rooted in a meaningful evidence base about impact on learning and teaching. Indeed, Graham Gibbs' influential 2010 report² concluded that class contact hours are not a valid indicator of teaching quality; it is what happens in those hours that matters, in addition to the quality and quantity of independent study.

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?

As noted in response to Question 13 above, Universities Scotland does not believe there is a credible way to construct such a measure.

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.

As noted in response to Question 13 above, Universities Scotland does not believe there is a credible way to construct such a measure.

The options presented in the consultation document are not realistically workable, nor are they rooted in a meaningful evidence base about impact on learning and teaching.

The particular option of a student survey poses several issues. For example, individual perspectives may not always reflect the full picture of learning opportunities on offer; preferences about type of contact vary amongst students; most students only experience one institution, but may respond to a survey in light of perceptions of other institutions;

² https://www.heacademy.ac.uk/system/files/dimensions_of_quality.pdf



and high response rates would be required to ensure the robustness of results, particularly if data were to be broken down to subject level.

The other options are all based on institutional reporting. It is not clear how any of these options could provide a meaningful picture of independent learning, a critical element of degree study. These options would also involve a significant additional burden on institutions. We have concerns as well about the potential use of student data for unintended purposes.

It is unclear how any of the options presented could provide reliably comparable information on the quality of contact that students are offered, as opposed to quantifiable aspects of contact. This is a fundamental problem for a framework purportedly aiming to measure teaching excellence. In general, there is also a risk in these proposals of generating perverse incentives for institutions to focus on quantity over quality of contact.

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?

The Scottish sector has engaged actively and constructively throughout the development of the TEF, and we welcome this opportunity to continue to contribute to the process as the TEF continues to evolve.

Although the sector has significant concerns about the present proposals, we remain committed to working with colleagues across the UK to explore options for a fit-for-purpose subject-level approach.

We also look forward to offering input to the upcoming Independent Review from a Scottish sector perspective.

