

TEF metrics peer review group: summary of activity

1. In 2019, the Office for Students (OfS) established a peer review group to undertake expert peer review of the OfS's proposed statistical methods for the future metrics to be used in the Teaching Excellence Framework (TEF). The TEF metrics peer review group (MPRG) were also tasked with commenting on the presentation and clarity of the new TEF metrics from a user perspective.
2. The MPRG includes members with expertise on the statistical elements of the TEF, and those who represent key metrics user groups. The group met 11 times between October 2019 and October 2021, and will operate up to the point of delivery of the new TEF framework later in 2022.¹
3. While the MPRG's remit is purely advisory, its members have played an important role in providing early advice on some of the proposals that we are now consulting on.² Their contributions have been acknowledged in the consultations on regulating student outcomes, the TEF and the construction of student outcome and experience measures for use in OfS regulation.³
4. This statement briefly summarises the topics we have discussed with the MPRG to date, and advice given by the group. We would like to thank the members of the group for their expertise, insight and challenge, noting that their advice does not constitute endorsement or otherwise of our proposals.

The MPRG's discussion topics

5. The group's discussions have covered a variety of topics, including the following:
 - Recommendations made by the Office for National Statistics (ONS)
 - Presentation of statistical uncertainty
 - Interpretation of indicator performance
 - Accommodating geographical influences in progression indicators
 - Approach to benchmarking, covering the review of the selection and grouping of benchmarking factors and adjustments to the underlying benchmarking calculations
 - Construction of the student experience and compound completion indicators.

¹ We are currently consulting on a timeline for the implementation of the new TEF framework, and have proposed that the next TEF assessments will commence from autumn 2022.

² Note, that in our consultation proposals we no longer refer to the term 'metrics' but instead refer to TEF indicators.

³ See www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/.

6. The group's discussions on each of the topics listed in paragraph 5 are summarised in paragraphs 7 to 29 below. These summaries are not intended to represent a comprehensive account of the group's consideration of each issue, several of which involved detailed discussions over a series of meetings. Instead, they draw out the key points which have influenced the further development of our proposals.

Recommendations made by the ONS

7. Throughout their work, the MPRG were mindful of recommendations made by the ONS in their review of the statistical elements of the TEF, on behalf of the independent review of the TEF.⁴ The independent review identified as particular priorities those ONS recommendations which highlighted the importance of effective communication of statistical uncertainty, and of the communication about the measures and methods used to inform TEF assessments more generally. Where ONS recommendations remained relevant to the proposed approaches for the future of TEF assessments, the group helped us consider the effectiveness of the improvements we were proposing for the TEF's use of data and indicators.

Presentation of statistical uncertainty

8. The group provided advice on the statistical properties of various options for the presentation of statistical uncertainty for both the indicator, the benchmark and the difference from benchmark. They advised on our assessment of the statistical assumptions that would be required for different options, and the extent to which these might be appropriate for different uses and users of statistics in different contexts. The MPRG also provided feedback on the potential opportunities and challenges for different users when engaging with the different presentation options.
9. The MPRG's contributions were sought at a formative stage in relation to certain aspects of the 'shaded bars' that we have proposed through our consultations. This included advice on the statistical methods that were appropriate to underpin the construction of the bars, as well as commenting on how easily users might understand and interpret them for the purposes of informing TEF submissions and assessments.⁵ The shaded bars represent the statistical uncertainty associated with observed values, such as the difference from benchmark. We have refined the presentation and communication of the shaded bars, following initial advice from the group and engagement with users.
10. When providing advice on the shaded bars, the MPRG identified a key benefit of an approach which maximised the communication of uncertainty as the ability to learn information about cohorts of any size (even if the full extent of that learned information was that there was too much uncertainty to learn anything about the provider's performance on the measure). In doing so, they were able to focus their advice on how different users could be empowered to make use of the statistics for their own purposes according to their own priorities. This set the scene for the group's engagement on a range of issues. For example, it meant that the group engaged with criteria of data suitability and suppression in a more limited way (to note the need for criteria aimed at preventing data disclosure in breach of the general data protection

⁴ See www.gov.uk/government/publications/independent-review-of-tef-report.

⁵ See proposal 11 of our consultation about the construction of student outcome and experience measures, with further statistical detail of its implementation described in the statistical methods supporting document. Both are available at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

regulation and ensuring a survey population was representative), and instead focused on the implications of data suppression for users of the statistics.

Interpretation of indicator performance

11. In parallel with the development of presentations intended to maximise the communication of statistical uncertainty, the group provided advice on approaches that might be used to support interpretation of a provider's performance as represented through the proposed indicators. Recognising that consistency of interpretation and judgements would be key for use of the data to inform TEF assessments, the group discussed a range of issues including:
 - a. **The opportunities and challenges in developing a formulaic approach to aid interpretation of the indicators data, either singularly or collectively.** In doing so, the MPRG's advice focused on identifying approaches that would maximise users' understanding of statistical uncertainty and empower TEF panel members to determine the strength of the evidence presented by the indicators, so that they could give this appropriate weight in reaching a judgement based on the totality of the available evidence (from indicators, the provider submission and potentially the student submission). Group members were supportive of approaches which would avoid the implementation of arbitrary weightings or thresholds in ways that could introduce 'cliff edge' effects. This includes minimising the reliance on statistical testing and the use of pre-determined significance levels (such as p-values).
 - b. **The visual aids that could be incorporated into presentations of the indicators to support consistency of interpretation.** In doing so, the MPRG considered approaches that could be appropriate specifically for TEF panel uses in line with any parameters defined through TEF assessment guidance. Their advice also considered the importance of maintaining the effective communication of statistical uncertainty to those users who might wish to apply different parameters to their interpretation of the indicators data.
 - c. **The practical implementation of our preferred presentations through interactive data dashboards, and priorities for the layout and navigation of these.** In doing so, the MPRG's advice focused on delivering a high-quality user experience within which users remained empowered to make use of the statistics for their own purposes according to their own priorities. The group considered the presentation of various aspects of the data dashboards (including detailed points of formatting) and feedback from different stages of our user testing programme (which involved OfS staff, members of the TEF advisory group and the OfS student panel).
12. The MPRG's contributions were sought at a formative stage in relation to proposals that the presentation of the data would include the following features:
 - a. Guiding lines overlaid onto the display of the shaded bars, which indicate where TEF assessment guidance will steer panel members to consider performance as materially above or below benchmark.⁶ While the group commented on the potential utility of this feature in conceptual terms, they agreed that the thresholds to be used were beyond

⁶ See proposal 11 and further guidance on interpretation available in Annex F of our consultation on the TEF, at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/the-tef/.

their remit and advised that any threshold should not be established solely on a statistical basis.⁷

- b. A summary table reported alongside the shaded bars to indicate the proportion of the uncertainty distribution represented by the bar that falls above or below the guiding lines, and the statistical calculations that would represent this most appropriately.⁸

Accommodating geographical influences in progression indicators

13. Having recognised the importance of contextualising measures of graduate outcomes and progression to recognise regional differences in labour markets across the UK, the MPRG provided early advice on the OfS's development of approaches for doing this. The group were mindful of the priority that the independent review of the TEF had placed on mechanisms for controlling for geographical considerations when interpreting and assessing measures of progression into employment or study.⁹ Their advice was tailored to this anticipated use and opportunities to account for these factors through the established benchmarking approach.
14. The MPRG's contributions were sought at various stages in the development of our experimental official statistics on a geography of employment and earnings.¹⁰ Our consultations propose use of the geography of employment quintiles described in those experimental official statistics, in the construction of split indicators and as a benchmarking factor for progression measures.¹¹

Approach to benchmarking

15. Our consultations propose the use of benchmarking to inform the assessments of a higher education provider's performance in respect of the student outcomes and experiences it delivers. The specific purpose and uses of benchmarking to inform OfS regulation of student outcomes and the TEF are explained in the relevant consultation documents.¹² In each case, benchmarks are used to help interpret a provider's actual performance relative to that of the sector overall, once we have taken into account the mix of students at the provider or the provision being offered, based on a set of benchmarking factors.

The review of the selection and grouping of benchmarking factors

16. The OfS has reviewed the selection and grouping of benchmarking factors to be used in the construction of different student outcome and experience measures and their benchmarks, with the work informing proposals included in Proposal 10 of the indicators consultation. The review

⁷ Our analysis that informed the proposed thresholds for these guiding lines are describing in the supporting document, 'Materiality and high benchmark values for use in interpretation of the TEF indicators', at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/the-tef/.

⁸ See proposal 11 and further guidance on interpretation available in Annex F of our consultation on the TEF, at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/the-tef/. Further statistical detail of its implementation are described in the statistical methods supporting document, at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

⁹ See www.gov.uk/government/publications/independent-review-of-tef-report.

¹⁰ See www.officeforstudents.org.uk/publications/a-geography-of-employment-and-earnings/.

¹¹ See proposals 9 (split indicators) and 10 (benchmarking factors) of our consultation about the construction of student outcome and experience measures, at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

¹² See www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/.

has sought to ensure that the resulting benchmarks are fit for purpose, based on the most recent evidence and analysis available.

17. The MPRG supported this work by providing advice on the formulation and application of the guiding principles that our consultation proposes for the selection and grouping of benchmarking factors.
18. The group also gave advice on the specifications of the detailed statistical modelling that informs the benchmarking review and identifies the correlation of a factor to the outcome being measured. However, their contribution did not extend to critiques of the judgement that the OfS has used to establish the best fit of different factors to the guiding principles.¹³

Adjustments to the underlying benchmarking calculation

19. Aspects of our proposed approach to the construction and reporting of student outcome and experience measures have required careful consideration of the ways in which our overall approach is reflected within the calculation of benchmarks and the application of their underlying statistical definitions. The MPRG supported this work by providing advice on the formulation and application of the statistical methods used.

The 'taught or registered' student population

20. To calculate the benchmarks which correspond to student outcomes and experience measures reported in relation to the population of students who are either registered or taught at the provider in question, it was necessary to consider a series of adjustments to the underlying statistical calculations. The adjustments were deemed necessary to accommodate the scenario that individual students would contribute to the indicators and benchmarks calculated for more than one provider, if they were taught under subcontractual arrangements.
21. The group considered the implications of this scenario with regard to calculation of the standard deviation of the difference from benchmark, and provided advice on the validity, testing and onward communication of the adjustments proposed. In recognition of the key role of the standard deviations for indicating statistical uncertainty, the MPRG's contributions included consideration of the impact of adjustments and descriptions of the revised covariance assumptions, as now included in our supporting 'Description of statistical methods' document.¹⁴

Potential solutions to problems of self-benchmarking

22. In their review of the statistical elements of the TEF, the ONS considered potential solutions to the problems of self-benchmarking. In doing so, they identified one approach as deserving of further consideration. When calculating the benchmark for an individual provider, the students at that provider contribute to the sector averages underpinning the calculation. The approach suggested by the ONS – referred to in their review as 'studentisation' – is a process which

¹³ A summary of the review of the selection and grouping of benchmarking factors has been published alongside the consultation, and describes the statistical methodology used for this purpose. See the supporting 'Review of the selection and grouping of benchmarking factors' document available at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

¹⁴ The description of this adjustment is described in Annex C of the statistical methods supporting document, at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

would see the removal of the contribution of each provider from the calculation of its own benchmark. The MPRG commented on this suggested approach in broad terms.¹⁵

23. Engaging with the issue at a formative stage, the group provided advice on the potential advantages and disadvantages of a method that would remove each provider's own students from the calculation of its benchmark, including the adjustments that may prove necessary within the underlying statistical calculations to implement such an approach.
24. The group's advice focused on possible conflicts with the proposed principles for the selection and grouping of benchmarking factors, including the potential to benefit from a reduction to the risks of self-benchmarking and the consequences of significant increases to both the complexity of the method and number of assumptions it relies upon. When providing initial advice on the issue, the MPRG recognised that the ONS had proposed it as a potential solution for a problem that primarily impacts a small minority of providers who were likely to be sufficiently distinctive that any benchmarking approach would encounter limitations to its effectiveness. They also provided insights as to the consequences of the approach for the onward understanding and interpretation of the wider benchmarking approach from their perspectives as users of the resulting statistics.
25. We have not proposed to make any adjustments to the benchmarking calculations through the use of the approach suggested by the ONS, nor through other means.¹⁶ To facilitate an understanding of where self-benchmarking may present a material issue for a given provider, we have proposed to include information about the provider's own contribution to that benchmark alongside the TEF indicators.

Construction of the student experience and compound completion indicators

26. Our consultations propose the construction and use of various measures of student outcomes and experiences for the purposes of different aspects of OfS regulation. These measures include measures of completion outcomes, and of the student experience based on responses to the National Student Survey (NSS). The MPRG were invited to comment on the construction of the measures in broad terms.

Student experience indicators

27. The student experience measures that the OfS has proposed to construct are based on responses to NSS question scales, by aggregating responses to individual questions within the scale. The MPRG were invited to consider exploratory analysis that sought to understand how the dependences between individuals' responses within a scale are accounted for. The group's contributions to the scope and specification of the analysis helped inform our subsequent conclusion that the NSS scale-based measures are constructed via appropriate and proportionate means, making best use of the individual question response data and taking

¹⁵ See www.gov.uk/government/publications/independent-review-of-tef-report. The ONS recommended that further consideration be given to a studentisation approach, including its implications for other parts of the process and the possible impact on robustness because of small sample sizes.

¹⁶ Described in paragraph 427 of proposal 10 of our consultation about the construction of student outcome and experience measures, at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

sufficient account of variations in a single student's responses to the questions within each scale.¹⁷

Compound completion indicators

28. The compound completion measure that the OfS has proposed to construct differs in its construction and benchmarking to that of other indicators. The MPRG's contributions helped us consider the accessibility of our communication of these approaches, as enacted through our consultations and supporting documents.¹⁸

Other matters

29. The MPRG also provided comments and feedback in relation to a range of other matters, such as the range and functionality associated with 'split indicators', data reportability thresholds and data about the size and shape of a provider's provision. In each case, their advice sought to ensure that:

- a. The approaches used would empower users of the statistics to make accurate interpretations and informed judgements, maximising their understanding of statistical uncertainty.
- b. The presentations and mechanisms included to guide interpretation, and any associated guidance, were focused on supporting consistent judgements that did not rely on the use of arbitrary thresholds or single, pre-determined significance levels.
- c. The number and granularity of the data points calculated would be informative but manageable for the intended purposes.

¹⁷ See proposal 8 of our consultation about the construction of student outcome and experience measures, at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

¹⁸ See proposal 6 of our consultation about the construction of student outcome and experience measures, at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.