

Office for
Students



Teaching Excellence and Student Outcomes Framework:

Subject-level pilot guide

Reference OfS 2018.44

Enquiries to tef@officeforstudents.org.uk

Publication date 22 October 2018



Foreword

By Professor Janice Kay, CBE

TEF Subject-level pilot Chair



As academics we relish the opportunity to challenge and test ideas.

The Teaching Excellence and Student Outcomes Framework (TEF) has generated plenty of robust debate and I have observed two emerging, complementary narratives over the last year.

Firstly, TEF has settled into being an accepted and valuable feature of the UK higher education landscape. It is also leading and inspiring international practice to focus on identifying and valuing excellent teaching, learning and student outcomes.

The 'provider-level' assessments ran smoothly and I have been struck by how the exercise has developed a sense of maturity in just two years. Most importantly, when I speak to colleagues throughout the sector, I hear more and more about how TEF is continuing to focus minds on genuine and meaningful improvements for students.

Secondly, people are starting to look at TEF as a long-term venture into raising esteem for teaching, driving enhancement and ensuring that the higher education sector continues to meet the needs of prospective and current students, whatever their backgrounds. The sector and its students have taken ownership and responsibility for its development.

TEF must stand for far more than simply being a good, technically sound process. To succeed it has to recognise, reward and drive meaningful improvements across the sector and underpin good quality information for students. It is right that we ask, 'What should TEF look like in ten years' time?', as Universities UK has done.

It is essential that we rigorously test and encourage debate on key aspects which will inform the final shape of subject-level TEF, and that is why it has been a huge privilege to lead the first subject-level pilot.

We have made significant progress already. The first year of pilots showed us that for the significant majority of provision we have a workable way forward, drawing on the best aspects of the two models tested, to accommodate and celebrate the diversity of provision across the UK.

I am pleased to clearly see the influence of all the lessons learned from the pilot and the government consultation in this guidance, which marks the next step on that journey of collaborative development. In particular, I welcome the recognition that some important aspects remain exploratory, and that there is more to unpick, challenge, clarify and simplify as we ask ourselves how best to achieve subject-level TEF that will have the most value for students, universities and colleges, industry and employers.

Janice Kay

Contents

Foreword	2
Contents	3
Enquiries and further information	5
Introduction	6
Transition to subject-level TEF	6
Development of subject-level TEF	7
Changes to the duration of provider-level TEF awards	8
Purpose of the 2018-19 subject-level pilot	9
The model for provider-level and subject-level assessment	15
Pilot design and structure	15
The relationship between provider- and subject-level assessment	18
The assessment criteria and framework	19
Assessment framework overview	19
Criteria.....	22
Evidence and assessment.....	22
Ratings	23
The definition of subjects and scope of subject assessments	29
What is a subject?	29
Which subjects will be assessed through the pilot?	31
Interdisciplinary provision	32
Single-subject providers	33
Evidence used in the assessments	35
Contextual information.....	36
TEF metrics	37
Submissions	43
Student engagement with submissions.....	50
The assessment process	51
The pilot panels	52
How assessment is organised	54
Relationship between subject- and provider-level assessment	57
Moderation and consistency	58
Method of assessment	59
Consideration of contextual information.....	60
Step 1: Review the metrics to form an initial hypothesis	61
Step 1a: Starting point for the initial hypothesis	61
Step 1b: Metrics-based initial hypothesis	63
Step 2: Review of the submission.....	71
Step 3: Best-fit holistic judgement.....	73
Outcomes	73
Ratings and statements of findings	73
Communication and use of outcomes.....	74
Evaluation of the pilots and next steps.....	74
List of abbreviations	75

Annex A: TEF Subjects list – mapping	76
Annex B: TEF subject-level pilot – full metrics descriptions	80
Metrics based on the National Student Survey	80
Continuation metrics.....	83
Employment metrics	85
Supplementary degree attainment data (at provider level only)	90
Annex C: Additional information for participating providers and TEF student representatives	93
Timeline.....	93
Requesting changes to data or scope of assessment.....	94
Submission instructions	96
Annex D: Subject panel roles and key responsibilities	99
Teaching Excellence Framework: Guide to subject-level pilot data	
(See separate document at www.officeforstudents.org.uk/advice-and-guidance/teaching/piloting-tef-at-subject-level/ .)	

Enquiries and further information

1. For enquiries about:

- participation in provider-level TEF or the development of subject-level TEF contact TEF@officeforstudents.org.uk.
- TEF metrics and the TEF extranet contact TEFmetrics@officeforstudents.org.uk.
- TEF policy contact the Department for Education (DfE) tef.queries@education.gov.uk.

2. For further information please see:

- the OfS's TEF subject-level pilot webpages: www.officeforstudents.org.uk/advice-and-guidance/teaching/piloting-tef-at-subject-level/
- the DfE's TEF webpages: www.gov.uk/government/collections/teaching-excellence-framework.
- For news and updates on the TEF from the OfS, we are launching a JiscMail list, TEFNEWS. To subscribe or unsubscribe from this list, visit <https://www.jiscmail.ac.uk/TEFNEWS>. Subscribers to the list will be kept informed of any updates and news relating to TEF, including publications, consultations, events and opportunities to apply to be a panel member.

If you need this publication in an alternative format, please contact Philip Purser-Hallard, philip.purser-hallard@officeforstudents.org.uk, 0117 931 7339, to discuss your needs.

Introduction

3. This guide explains how the second year of the Teaching Excellence and Student Outcomes Framework (TEF) subject-level pilots will be conducted, and what is being tested and evaluated in order to develop full subject-level TEF for future years of assessment. This guide also explains how participating universities and colleges, and their students, should take part, and how the pilot panels should undertake the assessments.
4. Participation in the pilot exercise is voluntary. In September the Office for Students (OfS) asked for expressions of interest¹ to take part from all eligible UK universities and colleges, and has selected a representative sample of around 50 of these to take part. Participating higher education providers, their student representatives and pilot panels will receive training and further guidance and instructions throughout the course of the pilot.
5. To support the development of subject-level TEF, all providers that are eligible to take part in the subject-level pilots will be given access to their own subject-level pilot data by the OfS. Providers not taking part in the pilot who want to understand how their subject-level pilot data would be used in the pilot assessment should also read this guide. This is in the interests of transparency and to support preparations for full subject-level TEF across the sector.
6. Particular guidance related to the subject-pilot's assessment of provider-level provision will be indicated by a **mauve border in the left margin**.
7. Particular guidance related to the subject-pilot's assessment of subjects will be indicated by a **green border in the left margin**.

Transition to subject-level TEF

8. The TEF was introduced in 2016 by the government to:
 - better inform students' choices about what and where to study
 - raise esteem for teaching
 - recognise and reward better teaching
 - better meet the needs of employers, business, industry and the professions.
9. To date, TEF has operated at 'provider level', which means that it has assessed the general performance across a university or college, producing a single rating for the whole provider. Participation in provider level TEF has been voluntary. Over 330 universities and colleges across the UK have taken part so far, receiving a rating of Gold, Silver, Bronze, or a Provisional award. The awards are based on a holistic assessment against common criteria using a set of metrics and a 15-page provider submission.
10. The DfE in England has committed to introducing TEF at subject level, to provide more useful information to prospective students about the subject they are looking to study. We know that most students choose their subject first, and then choose between universities and colleges

¹ The call for participants was published at www.officeforstudents.org.uk/news-blog-and-events/news-and-blog/ofs-seeks-higher-education-providers-to-pilot-tef-at-subject-level-for-the-second-year/.

offering that subject. We are piloting TEF at subject-level to provide information for applicants at both subject level and provider level. Alongside the pilot we will test applicants' reactions to such information and how to present it in the most useful way, recognising that applicants are making decisions about courses, and that subject-level TEF operates at a subject classification level. More information about how subjects are defined, for the purpose of this exercise, can be found at paragraphs 68-98.

11. Subject-level TEF will further drive enhancement of teaching and learning by identifying and incentivising excellence within subjects as well as at each provider. The outcomes will provide more granular information about how the institution performs as a whole, and across its subjects, including where excellence lies but also where it can make improvements.

Development of subject-level TEF

12. Subject-level TEF is being developed over two academic years (2017-18 and 2018-19) before it will be implemented. The first subject-level pilot (which ran in academic year 2017-18, and is referred to in this document as 'the first pilot') tested two options for producing both provider-level and subject-level ratings in one exercise. A representative selection of 50 providers from across the UK higher education sector took part².
13. The first pilot ran alongside the government's subject-level TEF consultation, which put forward detailed proposals on the design of subject-level TEF. The government consultation response and findings from the research have been published³. A research project which explored the TEF and student choice was also undertaken for the Department for Education (DfE) by IFF Research⁴.
14. The design of the second subject pilot (referred to in this document as 'the second pilot' or 'this pilot') has been informed by the findings from the first pilot, the government consultation and the student research.
15. The second pilot will be carried out in academic year 2018-19 alongside, but independently from, the provider-level TEF Year Four exercise⁵. The final design of subject-level TEF will take account of the pilot findings and the outcomes of the statutory independent review of TEF established by the Higher Education and Research Act, 2017 (HERA)⁶.
16. Up to now, each TEF exercise has been completed within a single academic year. However, given the scale of the first full subject-level TEF exercise, it will be conducted across two academic years, 2019-20 and 2020-21, to enable it to produce robust outcomes. This will ensure there is additional time for providers to make submissions and for panels to conduct the assessments. We expect the application window to open in early 2020, and to publish the outcomes in spring 2021. This will also allow more time for the findings of the second pilot and

² The results of the first pilot are available at: www.officeforstudents.org.uk/publications/teaching-excellence-and-student-outcomes-framework-findings-from-the-first-subject-pilot-2017-18/.

³ See www.gov.uk/government/consultations/teaching-excellence-and-student-outcomes-framework-subject-level.

⁴ See www.gov.uk/government/publications/teaching-excellence-framework-and-informing-student-choice.

⁵ See: www.officeforstudents.org.uk/advice-and-guidance/teaching/tef-year-four/

⁶ See section 26 of HERA at www.legislation.gov.uk/ukpga/2017/29/section/26/enacted.

the independent review to be fully considered before moving to full implementation. Figure 1 summarises the TEF development timeline.

Changes to the duration of provider-level TEF awards

17. The final provider-level exercise with published outcomes⁷ (TEF Year Four) will take place in 2018-19 and will operate completely independently from the subject-level pilots. So that subject-level TEF produces comprehensive outcomes to inform student choice, the DfE has decided that published awards from provider-level TEF Years Two, Three and Four should no longer be valid when subject-level TEF awards are published in 2021. At that point, all awards from provider-level TEF will expire, and be replaced by awards made through the first full subject-level TEF exercise (these awards will be at both provider and subject levels). To achieve this, the duration of awards made from provider-level TEF (TEF Years Two, Three and Four) will be adjusted as follows:

- a. All Year Two and Year Three awards due to expire in 2020 will be extended by one year (to 2021).
- b. Year Four awards will last for two years (to 2021).

18. Participation in provider-level TEF remains voluntary. If a provider has an award that is due to expire in 2019 and wishes to retain an award until subject-level TEF awards are published in 2021, it will need to participate in TEF Year Four. The TEF Year Four procedural guidance⁸ provide more information about participation and duration of provider-level TEF awards, and providers will receive a personalised email explaining what the changes mean for any current awards.

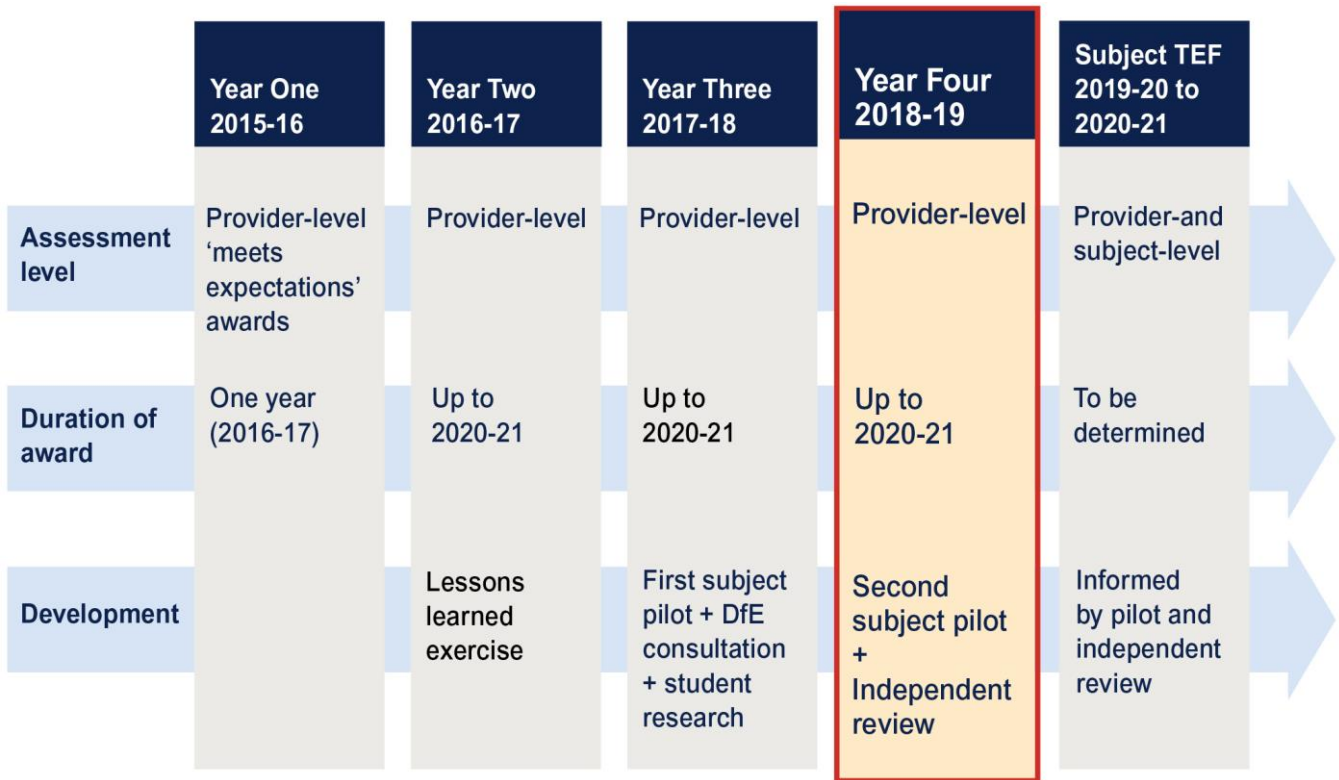
19. All eligible English providers with over 500 students⁹ will be required to take part in the first full subject-level TEF as a condition of registration with the OfS. Participation in full subject-level TEF will remain voluntary for other providers; these providers will need to participate if they wish to retain a TEF award beyond 2021. Figure 1 shows the timeline of TEF development.

⁷ See www.officeforstudents.org.uk/advice-and-guidance/teaching/tef-outcomes/#/.

⁸ Further detail about how provider-level awards will be adjusted to expire in 2021 is set out in the Year Four guidance, available at: www.officeforstudents.org.uk/advice-and-guidance/teaching/tef-year-four/.

⁹ The OfS is undertaking a consultation on student numbers: www.officeforstudents.org.uk/publications/calculating-student-numbers-consultation-on-the-principles-guiding-how-we-will-calculate-student-numbers-for-regulatory-purposes/.

Figure 1: Timeline of TEF development



Responsibilities

20. The OfS was established by the Higher Education and Research Act 2017 (HERA) and became fully operational in April 2018. The DfE commenced section 25 of HERA on 1 January 2018, which gives the OfS the power to make arrangements for a scheme to give ratings. The TEF is the scheme that the OfS has adopted under section 25 of HERA.¹⁰
21. The OfS has been responsible for implementing provider-level TEF according to the DfE's TEF specification, and has been working with the DfE to develop subject-level TEF. The OfS will carry out the second pilot collaboratively with participants and develop proposals for the final design of subject-level TEF. The DfE will consider the findings of the independent review of TEF and take policy decisions that will guide the final design of subject-level TEF.
22. A key feature of the TEF is that decisions about the ratings are made by independent panels at provider and subject level, comprising academics, students, employers and professional, statutory and regulatory bodies (PSRB) representatives. In the pilot, the pilot Main Panel will determine provider-level ratings; the pilot subject panels will determine subject-level ratings.

Purpose of the 2018-19 subject-level pilot

23. Following the first pilot and the government consultation, the second pilot will test a revised model of subject-level assessment, which combines successful features from the two models

¹⁰ See 'Securing student success: Regulatory Framework for higher education in England' (OfS 2019.01), available at www.officeforstudents.org.uk/publications/securing-student-success-regulatory-framework-for-higher-education-in-england/.

tested in the first pilot¹¹. It will also test and develop a series of other refinements to improve the meaningfulness of subject-level TEF to students, strengthen student engagement with submissions and assessment, test revisions to the evidence base and assessment method, and explore how to address limitations with the data at subject level. A summary of changes from the first pilot that are being tested in the second pilot is presented in Table 1.

¹¹ See www.gov.uk/government/consultations/teaching-excellence-and-student-outcomes-framework-subject-level. For details about the two models explored in the first pilot, see www.gov.uk/government/publications/teaching-excellence-framework-subject-level-pilot-specification.

Table 1: Summary of key refinements for the second pilot

Element	Summary of change
Subject classification system	The Higher Education Statistics Agency (HESA) Common Aggregation Hierarchy at level 2 (CAH2) will continue to be used with some refinements.
TEF framework	<p>Refinements to the criteria:</p> <ul style="list-style-type: none"> • separate the TEF criterion ‘TQ1: Student engagement’ into two distinct criteria • distinct criteria and rating descriptors at each level. <p>Refinements to the evidence:</p> <ul style="list-style-type: none"> • expand the range of contextual information • introduce two new National Student survey (NSS) metrics on learning resources and student voice (note that we will maintain the half-weighting of NSS metrics and the overall balance of NSS in the core metrics) • test a different combination of core metrics relating to student outcomes, drawn from Longitudinal Education Outcomes (LEO) and Destinations of Leavers from Higher Education (DLHE) data • test the feasibility of new data on differential degree attainment, and contextual data on grade inflation • test more directive guidance about the content of submissions • test an approach to ensuring that students have meaningful opportunities for involvement in submissions • test the inclusion of a provider summary statement to be used in subject-level assessment • revised examples of evidence that can be used in provider and subject-level submissions. <p>Refinements to the assessment and outcomes:</p> <ul style="list-style-type: none"> • revised formula for the starting point of the initial hypothesis • revised panel structure and processes, to ensure they will be scalable to full subject-level TEF • revisions to panel membership and roles, to enhance student membership on the pilot Main Panel and deploy widening participation (WP) expertise across subject panels • test how to make the outcomes more informative and useful for enhancement, including statements of findings at subject level.

Element	Summary of change
Model of assessment	<p>Comprehensive assessment of all subjects in a revised model that includes:</p> <ul style="list-style-type: none"> • provider-level assessment following a similar model to the current provider-level TEF • subject-level assessment with contextual information, metrics, submissions and ratings for each CAH2 subject • distinct criteria and rating descriptors at each level. <p>We will also test how the assessments at each level should interact.</p>
Limitations of data at subject level	<p>We will explore ways of mitigating limitations in the data at subject level, including:</p> <ul style="list-style-type: none"> • the application of minimum assessment thresholds for cohort sizes and number of data sources • approaches to maximising the use of available data • how submissions can best address data limitations • presentation of subjects that are not able to receive ratings.
Interdisciplinarity	<p>Will explore how far additional measures taken better accommodate interdisciplinary provision in subject-level assessment:</p> <ul style="list-style-type: none"> • students will continue to be counted pro rata in the subject-level metrics against each subject that their course is mapped to for all interdisciplinary provision • better contextual data at subject level that includes information about interdisciplinary provision will be used • specialist interdisciplinary panel member roles will be tested.

24. The purpose of the second pilot is to trial these refinements and fully test and evaluate the revised model. As the government response states, a ‘first priority should be to develop a robust model of assessment that produced meaningful ratings for students’. The final design of subject-level TEF that is taken through to full implementation will take account of the findings of the second pilot, as well as the outcomes of the independent review of the TEF.

25. We will test and evaluate the revised model against the following interacting themes:

- a. **Meaningfulness for students** – how the exercise engages students and whether the model generates subject-level ratings that are more meaningful and useful for students than the outputs of provider-level TEF alone.
- b. **Driving enhancement** – how the model incentivises and drives a focus on enhancement and improvements to learning and teaching and student outcomes for all students.

- c. **Supporting diversity of provision** – the capability of the model to recognise diverse and innovative forms of excellence across a diverse sector.
- d. **Supporting widening participation and social mobility** – how the model encourages providers to deliver positive outcomes for students from all backgrounds.
- e. **Robust evidence and assessment processes** – how well the model allows panels to make robust assessments, including how the metrics and submissions are used.
- f. **Effects on provider behaviour** – including the extent to which it avoids driving unintended consequences and minimises vulnerability to gaming.
- g. **Value for money** – the proportionality of cost of participation for providers and cost of delivery for government given the benefits for prospective and current students.

Meaningfulness for students

26. A key theme throughout this pilot will be to identify, explore and evaluate ways of making subject-level TEF more meaningful to students and strengthening their engagement in the process. As a result of student feedback gathered through the first pilot, we will explore how subject-level TEF can be best designed to support ongoing partnership between students and their universities and colleges to improve teaching quality and student outcomes.
27. We will seek to strengthen student involvement in both provider and subject-level submissions, and pilot evaluation activity through:
- a. Setting clear expectations for providers around involvement of students as full partners in considering and producing evidence for TEF assessments (see paragraphs 176-185).
 - b. Improving opportunities, and the support available to students to enable a full and meaningful contribution to submissions and evaluation activity, in part by the OfS establishing direct contact with and providing support for student representatives at participating providers.
 - c. Testing mechanisms for students to confirm their involvement (see paragraphs 178-185).
28. We will also test refinements to the assessment framework, as follows:
- a. Adapting the criteria to focus on what matters most to students. In particular, the TQ1 ‘Student engagement’ criterion will be split into two criteria: student engagement with learning, and student partnership.
 - b. Refining the basket of metrics to reflect what matters most to students. Specifically, two additional NSS question scales will be piloted as core metrics: Learning Resources, and the Student Voice.
29. Alongside the pilot, we will commission user testing of TEF outcomes with applicants. This activity will ensure that the content and presentation of TEF outcomes can support informed applicant choice alongside other information sources.

Driving enhancement

30. A key aim of TEF is to drive enhancement of teaching and learning. The pilot will seek to test ways in which subject-level TEF can achieve this. In particular, it will test producing statements of findings at both provider and subject level alongside the ratings, and explore the feasibility and usefulness of producing reports which identify and encourage areas of best disciplinary practice.

Supporting diversity of provision

31. The subject-level pilot will explore how far the revised model accommodates all types of providers' and provision, and evaluate whether measures to ensure rigour and consistency across assessments have been appropriately balanced against the need to accommodate diverse providers and forms of provision across the UK.

32. We will conduct the pilot with a representative and diverse selection of around 50 providers and their students, and have also sought diverse representation on the pilot panels.

Supporting widening participation and social mobility

33. The OfS is consulting on its approach to access and participation in higher education¹². We wish to develop a new strategy in which all the activities of the OfS work effectively together to improve equality of opportunity in student access, success and progression for groups of students who are currently underrepresented in English higher education, and the outcomes of this consultation will be considered in the design of the full subject-level TEF. The consultation places a stronger focus on reducing the gaps in success and progression, and we will in parallel develop and test how supplementary data based on differential degree attainment can be used within TEF.

34. The second pilot will increase the focus on how far the aspirations and achievements of students from all backgrounds are supported and achieved. We will test ways of enhancing the focus on differential outcomes between student groups through the metrics data, and by encouraging providers to identify and address differences in the experience and outcomes achieved by specific groups of students in their submissions. We will also explore how best to ensure WP expertise informs assessments across the subject panels.

Robust evidence and assessment processes

35. As we develop and prepare for full implementation of subject-level TEF we will focus on ensuring that assessments are technically robust, data limitations at subject level are addressed, and processes are designed in a way that is replicable and scalable.

36. Through the pilot we will test the revised model, refinements to the metrics and submissions, approaches to addressing data limitations, and a scalable panel structure and assessment processes. We will work with providers and panel members to develop and refine some of these aspects, and will seek their feedback on the robustness of the pilot assessments and which approaches to adopt or revise for full subject-level TEF.

¹² See www.officeforstudents.org.uk/advice-and-guidance/promoting-equal-opportunities/access-and-participation-plans/.

Effects on provider behaviour

37. The pilot will seek to explore the extent to which the design of subject-level TEF can avoid negative unintended consequences and minimise vulnerability to gaming.

Value for money

38. The government response notes that the revised model will lead to a higher burden on providers and a greater cost of running the exercise than the models tested in the first pilot. However it states that ‘the first priority should be to develop a robust model of assessment that produces meaningful ratings for students’.

39. Throughout the pilot we will work with providers and panel members to identify and fully explore whether any features of the revised model could be adjusted to reduce burden without compromising the principle of comprehensive assessment of all subjects. We will also consider burden and efficiency when designing and testing the submission and assessment processes.

The model for provider-level and subject-level assessment

Pilot design and structure

40. The first pilot demonstrated that the existing provider-level TEF framework (see Figure 2) and approach to assessment can be successfully extended to subject-level assessments. The model that we are testing will therefore continue to be based around the following features at both provider and subject level:

- independent peer review involving academics, students and other experts
- holistic judgements against a broad set of criteria, using a combination of quantitative and qualitative evidence, set within the context of the provider and its students
- meaningful and embedded student engagement throughout the process.

Figure 2: The high-level TEF framework

The high-level TEF framework	
Reference points , comprising three aspects of quality, criteria and rating descriptors.	Evidence , comprising: <ul style="list-style-type: none">• common data (including contextual data and TEF metrics) supplied by the OfS• provider submissions. Assessments , based on independent peer-review. Outcomes , comprising a statement of findings and TEF rating.

41. The revised model that will be tested in this pilot (demonstrated in Figure 3) consolidates the best features of the two models tested through the first pilot, whilst seeking to address issues it identified. The model has been conceptually tested through a series of workshops conducted with providers and panel members, including students, who were involved in the first pilot.

42. The revised model is comprehensive, in that it assesses all provision in the university or college. Each subject is assessed, and in parallel so is the university or college as a whole. The assessment process, evidence base, criteria and ratings descriptors are all tailored at subject and provider levels. The process for this model is shown in Figures 3 and 4.

Figure 3: Assessment model for the second pilot

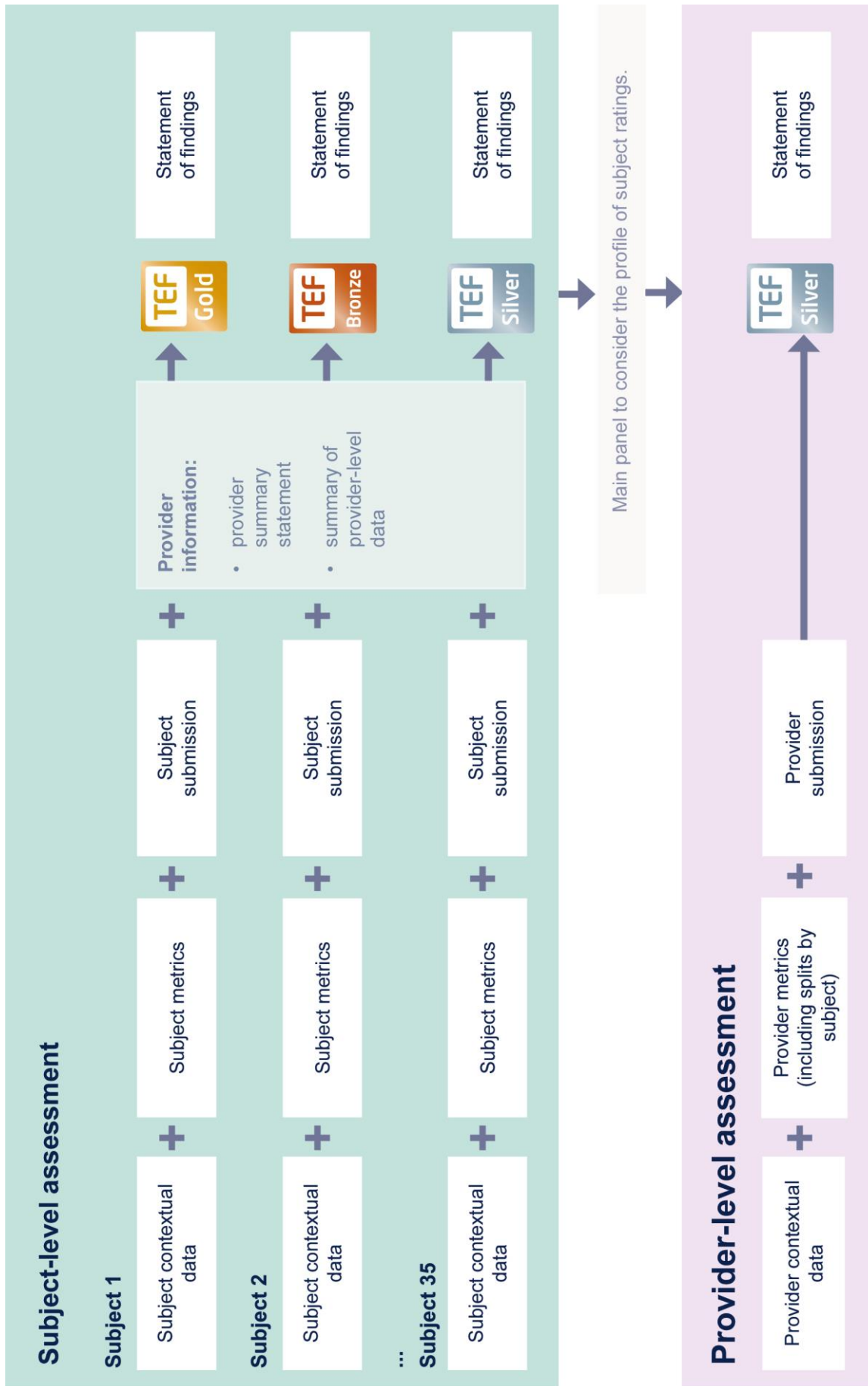
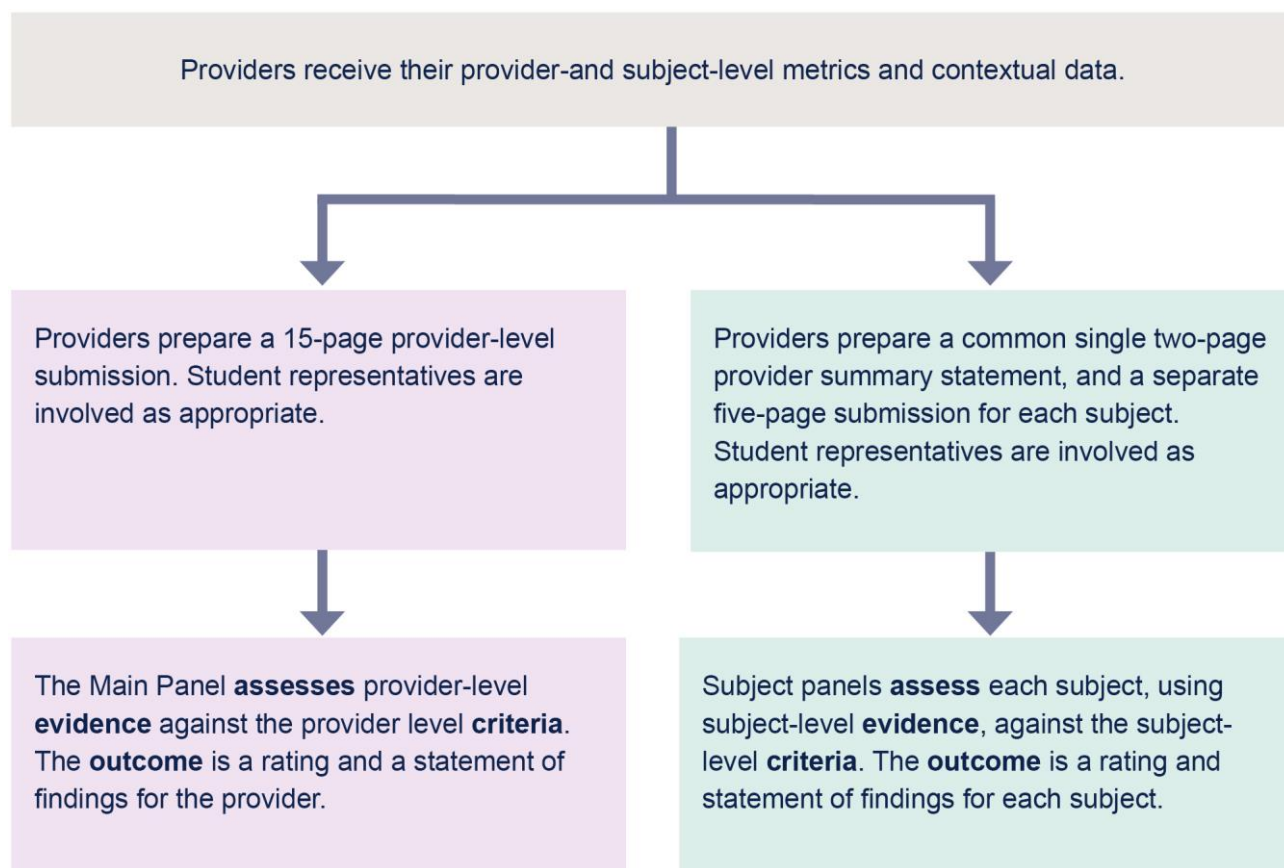


Figure 4: Assessment process for the second subject pilot



The relationship between provider- and subject-level assessment

43. Provider and subject-level assessments will be conducted separately from each other:

- a. Provider-level assessments will be undertaken by the pilot Main Panel and will be primarily based on provider-level evidence, assessed against the provider-level criteria.
- b. Subject-level assessments will be undertaken by subject panels and will be primarily based on subject-level evidence, assessed against the subject-level criteria.

44. Although the assessments will be conducted at each level separately, we will test some linkages between the two and seek to achieve an appropriate coherence between them. This recognises that teaching and learning in a subject takes place in the context of the provider; and that the overall performance of a provider is reflected at least in part by teaching and learning across its subjects.

45. In particular, we will test the following linkages:

- a. Using a common set of 11 criteria which are tailored at provider and subject levels.
- b. The provider-level assessment will include some information broken down into subjects:
 - i. The metrics at provider-level will include splits of performance by subject, so that the Main Panel can consider consistency across subjects as well as student groups.

- ii. There will be an opportunity for the Main Panel to review provider-level ratings in light of the final profile of the subject-level ratings that have been awarded by subject panels.
- c. The subject-level assessment will include some information about the whole provider:
 - i. After initially considering information purely at subject level, the subject panels will also consider information about the whole provider and how this affects performance in the subject. We will test how far this is useful and relevant at subject level.
- d. There will be interaction between the pilot Main Panel and subject panels:
 - i. The chairs and deputies of each subject panel are also members of the Main Panel.
 - ii. The Main Panel will provide oversight of the subject panels to ensure consistency in their approach to assessment across subjects, and ensure coherence between subject and the provider-level assessment.

46. In parallel to conducting the subject pilots we will test applicants' reactions to provider- and subject-level outcomes, and test possible approaches to presenting them, to best support applicant choice.

The assessment criteria and framework

Assessment framework overview

47. Teaching quality is best considered in the context of students' learning. The outcomes of students' learning are determined by the quality of teaching they experience, the additional support for learning that is available, and what the students themselves put into their studies. This should be supported, facilitated and delivered by the provider and its subjects.
48. The TEF has been designed, in partnership with sector representatives and sector agencies, to focus on three aspects of quality:
- teaching quality
 - learning environment
 - student outcomes and learning gain.
49. Providers and subjects will be assessed in relation to these three aspects, which are broken down into 11 common criteria. Table 2 summarises the criteria and associated evidence related to each aspect. The assessment is holistic and panels will make a single overall judgement based on the totality of all the evidence, across all three aspects.
50. The overall framework applies to both provider and subject level, though there are some detailed differences in terms of:
- the way the criteria and ratings are described at provider and subject level, to recognise the different expectations and responsibilities at each level
 - the details of some of the contextual and metrics data

- the page length for submissions
- the information considered in each step of the assessment.

Table 2: TEF assessment framework

		Provider level	Subject level
Criteria		Criteria defined at provider level:	Criteria defined at subject level:
		<p style="text-align: center;">Teaching quality (TQ):</p> <ul style="list-style-type: none"> • Student engagement with learning (TQ1) <ul style="list-style-type: none"> • Valuing teaching (TQ2) • Rigour and stretch (TQ3) <ul style="list-style-type: none"> • Feedback (TQ4) • Student partnership (TQ5) <p style="text-align: center;">Learning environment (LE):</p> <ul style="list-style-type: none"> • Resources (LE1) <ul style="list-style-type: none"> • Scholarship, research and professional practice (LE2) <ul style="list-style-type: none"> • Personalised learning (LE3) <p style="text-align: center;">Student outcomes and learning gain (SO):</p> <ul style="list-style-type: none"> • Employability and transferable skills (SO1) • Employment and further study (SO2) <ul style="list-style-type: none"> • Positive outcomes for all (SO3) 	
Evidence	Contextual information	<ul style="list-style-type: none"> • Provider-level contextual data • Maps of provider-level geographic context 	<ul style="list-style-type: none"> • Subject-level contextual data • Maps of subject-level geographic context
	Metrics	Provider-level metrics:	Subject-level metrics:
		<ul style="list-style-type: none"> • Teaching on my course (NSS) • Assessment and feedback (NSS) <ul style="list-style-type: none"> • Student voice (NSS) • Academic support (NSS) • Learning resources (NSS) • Continuation (HESA/ILR) • Highly-skilled employment or higher study (DHLE) • Sustained employment or further study (LEO) • Above median earnings threshold or higher study (LEO) 	
	Supplementary data	Where applicable: <ul style="list-style-type: none"> • Differential degree attainment data and grade inflation data • Additional data on part-time provision 	
Submissions	<ul style="list-style-type: none"> • Provider-level submission (up to 15 pages) 	<ul style="list-style-type: none"> • Subject-level submission (up to five pages) • Provider summary statement (up to two pages) 	
Assessment		Assessment by the pilot main panel: <ul style="list-style-type: none"> • Three-step method of assessment • Best-fit judgement against the provider-level rating descriptors 	Assessment by the pilot subject panels: <ul style="list-style-type: none"> • Three-step method of assessment • Best-fit judgement against the subject-level rating descriptors
Outcomes		Provider-level rating and statement of findings	Subject-level rating and statement of findings

Criteria

51. There are 11 criteria mapped against the three aspects of quality. Each criterion is described at both provider and subject level. The criteria have been designed to:

- focus on the issues that contribute and matter most to student experiences of teaching, learning and student outcomes
- positively articulate features of excellence in relation to these issues, which providers and subjects may meet to an extent, or exceed
- recognise diverse forms of excellence and avoid constraining innovation.

52. Each provider will have a distinct approach to managing teaching excellence and achieving excellent student outcomes in relation to the criteria. This means that the way resources and responsibilities are allocated and managed across the provider, its services and its subjects vary from one provider to the next according to what works best for providers and their students. TEF assessments will focus only on the impact of different approaches on the student experience and outcomes, rather than seeking to make judgements on intentions, activity, or strategy which have not yet demonstrated an impact.

53. The complete provider-level and subject-level criteria for the pilot can be found in Table 3. They have been revised and tailored at provider and subject level in consultation with participants and panel members from the first pilot. We expect to further refine them, and the rating descriptors, through the second pilot.

54. At provider level, the criteria focus on:

- institutional policy and strategy
- an enabling or facilitating institutional environment
- provider-wide services, facilities and resources
- the relationship with subject-level provision
- outcomes for all groups of students.

55. At subject level, the criteria focus on:

- curriculum, pedagogy and practice
- the experience and outcomes of students studying the subject
- the embedding of institutional policy and strategy in the subject.

Evidence and assessment

56. The evidence used to assess each provider and to assess each subject will comprise:

- a standard set of metrics, produced by the OfS
- a submission made by the provider, with appropriate student involvement.

57. Independent panels of experts will carry out the assessments, applying their judgement to the combination of metrics and submitted evidence, against the criteria. Contextual information is also provided at provider and subject levels to inform and orientate panel members for assessment.
58. The pilot Main Panel will be collectively responsible for deciding the provider-level ratings.
59. Each pilot subject panel will be collectively responsible for deciding the subject-level ratings for all subjects within its remit.
60. The Main Panel will provide oversight, and we will test mechanisms for the Main Panel to ensure consistent standards of TEF assessment across the subject panels.
61. The panels will follow a structured method of assessment to maximise consistency in judgements and fairness of assessment; the final judgement, however, is holistic and takes into account all the evidence.
62. Specifically, each assessment will begin by considering standardised contextual information. Next, the panel members will form an initial hypothesis based on the metrics. They will then review the submission, considering how far it provides evidence of excellence in relation to all three aspects of quality. Having considered all the available evidence, the panel members make a holistic judgement about which of the rating descriptors is a 'best fit'. Ratings descriptors are given in Table 4.
63. More information about the evidence used can be found in paragraphs 99-185. Further information about how the assessments will take place – including how evidence will be considered at both provider and subject levels to form judgments – can be found in paragraphs 186-293.

Ratings

64. Provider-level assessment in the pilot will result in one of three possible indicative ratings: Bronze, Silver, or Gold.
65. Subject-level assessment in the pilot will also result in Bronze, Silver, Gold ratings, but we will also test circumstances where a subject may not receive a rating, where the panel deems there is insufficient evidence to make a 'best fit' judgement. In addition, some subjects will not be assessed where they are out of scope or do not meet minimum data requirements to be assessed. We will test the options for how to best present subjects that are not assessed or do not receive a rating.
66. Table 4 gives the full ratings descriptors for all awards. These have been tailored at provider and subject levels.
67. In addition to generating ratings, the pilot will test how statements of findings at both provider and subject levels can support enhancement and inform applicants.

Table 3: TEF criteria at provider and subject level

Teaching quality	Provider-level criteria are the extent to which:	Subject-level criteria are the extent to which:
Student engagement with learning (TQ1)	The provider leads enhancement of practices that encourage effective student engagement with learning.	Students are engaged with learning and encouraged to commit to their studies, including through appropriate contact time and independent learning.
Valuing teaching (TQ2)	The provider fosters a culture that promotes, recognises and rewards excellent teaching.	Excellent teaching in the subject is promoted and recognised, and innovation is supported.
Rigour and stretch (TQ3)	The provider maintains high standards that challenge students to develop independence, knowledge and skills that reflect their potential.	Curriculum, pedagogy and assessment stretch students to develop independence, knowledge and skills that reflect their potential.
Assessment and feedback (TQ4)	The provider's assessment and feedback policies effectively support students' development, progression and attainment.	Assessment and feedback practices effectively support students' development, progression and attainment.
Student partnership (TQ5)	The provider understands and works in partnership with its students to influence the strategic approach to, and enhancement of, provision.	Students studying the subject are understood and engaged as partners in the delivery, development and enhancement of provision.

Learning environment	Provider-level criteria are the extent to which:	Subject-level criteria are the extent to which:
Resources (LE1)	The provider's physical and digital resources support and enhance teaching and learning.	Physical and digital resources are used effectively to aid teaching, learning and the development of independent study and research skills.
Scholarship, research and professional practice (LE2)	The provider facilitates a learning environment that is enriched by scholarship, research and/or professional practice.	Teaching and learning is enriched by student exposure to and involvement in scholarship, research and professional practice.
Personalised learning (LE3)	The provider facilitates personalised learning and a supportive environment that enables student progression and attainment.	Students are supported as individuals to succeed in their learning, progression and attainment.

Student outcomes and learning gain	Provider-level criteria are the extent to which:	Subject-level criteria are the extent to which:
Employability and transferable skills (SO1)	The provider strategically enables its students to gain knowledge, skills, work-readiness and personal development that equip them for their careers, and enhance their personal and professional lives.	Students gain knowledge, skills, work-readiness and personal development, which are embedded within the curriculum and wider student experience.
Employment and further study (SO2)	The provider effectively supports students through their academic and wider experiences to progress into graduate level employment or higher study.	Students progress educationally and professionally, into graduate level employment or higher study.
Positive outcomes for all (SO3)	The provider's students from all backgrounds achieve positive outcomes. The provider strategically identifies and addresses potential differential outcomes of disadvantaged students	Students from all backgrounds achieve positive outcomes. Potential differential outcomes of disadvantaged students are identified and addressed.

Table 4: Ratings descriptors of Gold, Silver, Bronze TEF awards for both provider and subject level



The provider supports consistently outstanding teaching, learning and student outcomes for its students. It is of the highest quality found in the UK higher education sector.

The TEF panel has judged that the provider's performance across the three aspects of teaching quality best fits the descriptor below.

Teaching quality

The provider drives enhancement in practices that encourage outstanding student engagement with learning. There is an embedded culture that effectively promotes, recognises and rewards outstanding teaching across the provider. The provider rigorously maintains high standards that consistently challenge its students to achieve their full potential. Embedded assessment and feedback policies effectively support students' development, progression and attainment. The provider has proactive, embedded and effective ways of understanding and working in partnership with its students, who influence the strategic approach to, and enhancement of, provision.

Learning environment

Outstanding physical and digital resources support and enhance teaching and learning across the provider. The learning environment is enriched by cutting-edge scholarship, research and/or professional practice. The provider effectively supports personalised learning and its students achieve outstanding levels of progression and attainment.

Student outcomes and learning gain

The provider actively ensures its students from all backgrounds gain knowledge, skills, work-readiness and personal development that enhance their careers and personal lives. The provider strategically and effectively supports its students to progress beyond their studies, and they achieve outstanding levels of progression into graduate-level employment or higher study. The provider's students from all backgrounds achieve positive outcomes. Potential differential outcomes for disadvantaged students are strategically and effectively identified and addressed.



Teaching, learning and student outcomes in this subject are outstanding, and are of the highest quality found across the UK in this subject.

The TEF panel has judged that the subject's performance across the three aspects of teaching quality best fits the descriptor below.

Teaching quality

Students are highly engaged with and actively committed to their studies, including through appropriate contact time and independent learning. Outstanding teaching in the subject is promoted and recognised, and innovation is effectively supported. The curriculum, pedagogy and assessment stretch students to meet their full potential. Embedded assessment and feedback practices effectively support students' development, progression and attainment. Students in the subject are well understood and are actively and effectively engaged as partners in the delivery, development and enhancement of provision.

Learning environment

Outstanding physical and digital resources are used effectively to aid teaching, learning and the development of independent study and research skills. Teaching and learning in the subject is enriched by student exposure to and involvement in cutting-edge scholarship, research and/or professional practice. Students are effectively supported as individuals to succeed in their learning and achieve outstanding levels of progression and attainment.

Student outcomes and learning gain

Students gain knowledge, skills, work-readiness and personal development, which are effectively embedded and promoted in the curriculum and wider student experience. Students achieve outstanding levels of progression into graduate-level employment or higher study. Positive outcomes are achieved by students from all backgrounds. Potential differential outcomes for disadvantaged students are effectively identified and addressed.



The provider supports excellent teaching, learning and student outcomes for its students. It consistently exceeds the rigorous national quality requirements for the UK higher education sector.

The TEF panel has judged that the provider's performance across the three aspects of teaching quality best fits the descriptor below.

Teaching quality

The provider leads enhancement in practices that encourage high levels of student engagement with learning. There is a culture that promotes, recognises and rewards excellent teaching across the provider. The provider maintains high standards that challenge its students to achieve their potential. Assessment and feedback policies effectively support students' development, progression and attainment. The provider has effective ways of understanding its students and enabling students to influence the strategic approach to, and enhancement of, provision.

Learning environment

High-quality physical and digital resources support and enhance teaching and learning. The learning environment is enriched by scholarship, research and/or professional practice. The provider supports personalised learning and its students achieve high levels of progression and attainment.

Student outcomes and learning gain

The provider effectively enables its students to gain knowledge, skills, work-readiness and personal development that enhance their careers and personal lives. The provider effectively supports its students to progress beyond their studies, and they achieve high levels of progression into graduate-level employment or higher study. The provider enables students from all backgrounds to achieve positive outcomes. Potential differential outcomes for disadvantaged students are identified and addressed.



Teaching, learning and student outcomes in this subject are excellent.

The TEF panel has judged that the subject's performance across the three aspects of teaching quality best fits the descriptor below.

Teaching quality

Students are highly engaged with and encouraged to commit to their studies, including through appropriate contact time and independent learning. Excellent teaching in the subject is promoted and recognised, and innovation is supported. The curriculum, pedagogy and assessment encourage students to meet their potential. Assessment and feedback practices effectively support students' development, progression and attainment. Students in the subject are understood and engaged as partners in the delivery, development and enhancement of provision.

Learning environment

High-quality physical and digital resources are used effectively to aid teaching, learning and the development of independent study and research skills. Teaching and learning in the subject is enriched by student exposure to and involvement in scholarship, research and/or professional practice. A range of individual learning styles are supported and students achieve high levels of progression and attainment.

Student outcomes and learning gain

Students gain knowledge, skills, work-readiness and personal development, which are embedded within the curriculum and wider student experience. Students achieve high levels of progression into graduate-level employment or higher study. Students from all backgrounds are supported to achieve positive outcomes. Potential differential outcomes for disadvantaged students are identified and addressed.



The provider supports good teaching, learning and student outcomes for its students. It meets rigorous national quality requirements for the UK higher education sector.

The TEF panel has judged that the provider's performance across the three aspects of teaching quality best fits the descriptor below.

Teaching quality

The provider encourages and supports student engagement with learning. Good quality teaching across the provider is supported and recognised. The provider maintains high standards that enable its students to achieve their potential. Assessment and feedback policies support students' development, progression and attainment. The provider seeks to understand its students and considers their feedback as part of its strategic approach to, and enhancement of, provision.

Learning environment

Physical and digital resources support and enhance teaching and learning. The learning environment is enriched by some scholarship, research and/or professional practice. The provider supports its students' learning, progression and attainment.

Student outcomes and learning gain

The provider enables its students to gain knowledge, skills, work-readiness and personal development that enhance their careers and personal lives. The provider supports its students to progress beyond their studies into graduate-level employment or higher study. The provider enables students from most backgrounds to achieve positive outcomes, and seeks to identify and address potential differential outcomes of disadvantaged students.



Teaching, learning and student outcomes in this subject are good.

The TEF panel has judged that the subject's performance across the three aspects of teaching quality best fits the descriptors below.

Teaching quality

Students are engaged with their studies, including through appropriate contact time and independent learning. Good quality teaching in the subject is promoted and recognised. The curriculum, pedagogy and assessment enable students to meet their potential. Assessment and feedback practices support students' development, progression and attainment. Student feedback informs the delivery, development and enhancement of provision.

Learning environment

Physical and digital resources are used to aid teaching, learning and the development of independent study and research skills. Teaching and learning in the subject is enriched by some student exposure to and involvement in scholarship, research and/or professional practice. Students are supported in their learning, progression and attainment.

Student outcomes and learning gain

Students gain knowledge, skills, work-readiness and personal development, which are included in the curriculum and wider student experience. Students are supported to progress into graduate-level employment or higher study. Students from most backgrounds are supported to achieve positive outcomes.

The definition of subjects and scope of subject assessments

68. The scope of subject-level TEF covers all undergraduate provision leading to qualifications at the appropriate levels of the national frameworks for higher education qualifications, and all modes of delivery, including full-time, part-time, distance, work-based and blended learning.
69. To participate in the pilot, a provider must be eligible for TEF and have suitable metrics at provider level as defined for TEF Year Four¹³, (see ‘Guide to subject-level pilot data’¹⁴ paragraphs 112-114). We are not, as part of the pilot, testing criteria or processes for determining the eligibility of a provider to take part in TEF in future. These criteria and processes will be considered in the light of final decisions on the basket of metrics, and the implications of regulatory changes in England, in which participation in subject-level TEF will be a condition of registration for providers with more than 500 students.
70. For the providers participating in the pilot, we will test criteria and processes for determining which **subjects**:
- are ‘in scope’
 - have sufficient data for assessment at subject level (see paragraph 76 and Annex C, figure C1).

What is a subject?

71. All students on undergraduate courses are assigned, by their provider, to between one and three subject codes to reflect the subject or subjects their overall qualification relates to.
72. The detailed subject codes are aggregated according to the HESA Common Aggregation Hierarchy (CAH).
73. We have adapted level 2 of the CAH for the purpose of this pilot, based on feedback from the first pilot and responses to the DfE consultation on subject-level TEF. However, rather than create a bespoke aggregation we will work with HESA to review and align any changes following the pilot, to maintain a single CAH which can be used for multiple purposes, including the TEF. Table 5 shows the subject categories that will be used in this pilot, alongside the pilot subject panels that will assess the subjects. This is explained in paragraphs 188-199. Annex A shows how these subject categories are mapped to level 3 of the CAH, and shows which subjects are within the remit of each pilot subject panel.

¹³ See www.officeforstudents.org.uk/advice-and-guidance/teaching/tef-year-four/technical-guidance-for-participants/.

¹⁴ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

Table 5: Subject categories

Pilot Subject Panel	CAH2 subjects to be assessed (adapted for the TEF pilot)
Medical sciences +	Medicine and dentistry
	Pharmacology, toxicology and pharmacy
	Medical sciences
	Psychology
	Veterinary sciences
Nursing and allied health subjects	Nursing and midwifery
	Allied health
	Sport and exercise sciences
Natural sciences +	Biosciences
	Physics and astronomy
	Chemistry
	General, applied and forensic sciences
	Mathematical sciences
Engineering and technology	Engineering
	Materials and technology
	Computing
Natural and built environment +	Agriculture, food and related studies
	Geography, earth and environmental studies
	Architecture, building and planning
Social sciences	Sociology, social policy and anthropology
	Economics
	Politics
Business and law +	Law
	Business and management
Education and social care	Education and teaching
	Health and social care
Arts +	Creative arts and design
	Performing arts
	Media, journalism and communications
Humanities	English studies
	Languages and area studies
	History and archaeology
	Philosophy and religious studies
	Combined and general studies

74. Often a number of courses make up a particular 'subject' at a provider. Most students study on courses which clearly map into a particular subject group. However, where students study courses which involve more than one subject, the sum of their data will be split and

apportioned to more than one subject groups. For more information about this, see paragraph 81-90.

Which subjects will be assessed through the pilot?

75. All subjects at participating providers that are 'in scope' and have sufficient metrics data will be assessed in the pilot in their own right as separate subjects. We will test criteria and processes for determining these.
76. The first pilot identified some subjects at providers that were inappropriate to assess (for example, where all the provision in a subject had been discontinued). In the second pilot we will treat subjects made entirely of courses that are no longer recruiting students as out of scope for a subject-level assessment. We will test a process for filtering out such subjects. Further detail can be found in the 'Guide to subject-level pilot data', paragraphs 115-120.¹⁵ Participating providers will be provided by the OfS with a data-driven indication of their subjects that appear to be in and out of scope, based on whether they have recently recruited students in that subject. The provider will be able to review and request changes to this, if necessary (see Figure 1 in the 'Guide to subject-level pilot data'¹⁶).
77. The first subject-level pilot, and responses to the DfE consultation, also identified limitations in how far the data in some subjects could inform robust assessments. The second pilot will test ways of generating robust ratings for as wide a range of subjects as possible, while also setting thresholds for the minimum data required to inform a subject-level assessment. A provider's subjects that are in scope will be assessed where they have sufficient data for assessment, as follows:
 - a. At least two of the following TEF 'metric types' are reportable in the subject's majority mode of study: NSS-based metrics, continuation, or employment metrics (either the LEO or DLHE metrics).
 - b. The data covers at least 20 students for the subject's majority mode of study (as measured in the contextual data – see 'Guide to subject-level data', paragraphs 26-33¹⁷).
78. For the purpose of this pilot, the threshold of at least 20 students is deliberately lower than the first pilot indicated might be necessary to inform robust assessments. This is so that we can test the lower limits with a view to maximising the coverage of subject-level TEF. To help address limitations in the data relating to small cohorts we will:
 - a. Test with the panels whether they can make more use of available metrics data.
 - b. Work with the pilot providers and panels to test whether submissions might be able to mitigate data limitations and enable robust assessments.
79. We will also test ways of presenting TEF outcomes for subjects that are out of scope or do not have assessable data, as part of the overall package of subject-level TEF outcomes for

¹⁵ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

¹⁶ Ibid.

¹⁷ Ibid.

a provider. Particular attention will be given to avoiding negative implications where a subject does not meet the requirements for assessable data.

80. Where a subject is not assessed (whether it is out of scope or does not have assessable data), the data for that subject will still be included in the provider-level data, and hence remain part of the provider-level assessment.

Interdisciplinary provision

81. A number of considerations have been taken into account regarding how interdisciplinary provision will be accommodated in this subject pilot.
82. 'Interdisciplinary provision' is a term that in itself can reflect a variety of different types of arrangements – for the return of associated student data as well as for courses offered to potential applicants and for the delivery of the provision. For the purposes of the subject assessment, we are primarily concerned with courses where metrics span multiple CAH2 subjects.
83. In some subjects, the advantages of an interdisciplinary approach may be reflected in metrics data and require no further explanation. Where providers wish to draw out the impact of their interdisciplinary approach separately from the TEF metrics, they will be able to do so through the relevant written submissions, for example in each of the individual subjects involved.

Multi-subject programmes

84. When we refer to 'multi-subject' courses we are discussing those which span multiple CAH2 subjects. The arrangements regarding this type of provision are sometimes clear to students and other audiences through the presentation and labelling of the course – for example a joint honours degree in history with French, or a joint maths with economics course – and in the reporting practices for the associated student data. There may, however, be cases where it is unclear from the course's title that it spans multiple CAH2 subjects – a biochemistry course, for example.
85. Courses can currently be assigned up to three Joint Academic Coding System (JACS) or Learn Direct Classification System codes by the provider, each of which may potentially map to separate CAH2 subjects within TEF. For example, the biochemistry course may be assigned to codes that identify biosciences, as well as codes that identify chemistry.
86. Each individual student on a course is assigned to the relevant CAH2 subjects, pro rata to the proportional course distribution identified by the provider in the student data return, with the outcome for that student apportioned to the metrics for those subjects on the same basis. If the course has been identified as 50 per cent each in biosciences and chemistry, this would mean that a biochemistry student who continues in higher education in the year after beginning their course counts positively towards the continuation metrics for each of the two subjects, weighted as 0.5 in the numerator and denominator for each of the subjects. More information about this can be found at in the 'Guide to subject-level pilot data', paragraphs 20-21.¹⁸

¹⁸ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

87. Ratings are awarded for discrete CAH2 subjects; separate ratings are not awarded for joint programmes.

Modular degrees and other interdisciplinary approaches

88. The interdisciplinary nature of provision may be conceptual, in the design of modular degrees and other very broadly based provision. Examples might include liberal arts or natural sciences degrees.

89. The assignment of students to CAH2 subject areas for TEF purposes uses the JACS or Learn Direct Classification System codes that a provider assigns to the course as a whole. The nature of the interdisciplinary provision we are discussing may not be evident at this level, and only visible in the codes that a provider assigns to the modules comprising the course. However, it is also possible for a provider to assign a course to a single, generic JACS code (Y000) which is reserved for use in rare, exceptional cases where a provider considers it unsuitable to assign the course to up to three alternative codes.

90. If the course has been identified against the Y000 JACS code, the student will be mapped to the 'combined and general studies' subject area. It follows that even for an individual provider, the provision identified as 'combined and general studies' may not comprise homogeneous groups of students experiencing higher education in a similar way. The benchmarking, which looks at students mapped to this subject across all providers in the sector, should be treated with particular caution. This means that the associated flags and step 1a metrics-based initial hypothesis must be weakly held, and panel members will be guided to place greater weight on the evidence provided in the submission. See paragraphs 231-284 for further information on how metrics and submissions should be considered to inform assessments.

Interdisciplinary expertise

91. In recognition of the complexity of interdisciplinary provision, in this pilot, each subject panel will have a member appointed as the interdisciplinary liaison.

92. The interdisciplinary liaison will be a subject panel member who has substantive experience of interdisciplinary teaching and learning. They will be able to advise other panel members in their discussions of cases concerning interdisciplinary provision, raise the panel's awareness of interdisciplinary provision, and liaise with other subject panels as appropriate. See paragraphs 188-199 for further information on panel structures.

93. We will want to test how far these approaches accommodate interdisciplinary provision, and will seek to identify further opportunities, and ways in which existing opportunities might be strengthened through the pilot activity.

Single-subject providers

94. In the first pilot, single-subject providers were identified as those whose provision was delivered entirely within one CAH2 subject. In the second pilot, the introduction of criteria tailored at provider and subject levels, and criteria for identifying subjects that are in scope and have sufficient data for assessment, introduces the scenarios given in Table 6.

Table 6: Single- and multi-subject providers

Provider type	Subject-level assessment	Metrics
Single-subject provider	The entirety of the provider’s provision is within a single CAH2 subject.	The subject- and provider-level metrics are identical.
Multi-subject provider that has only one assessable subject	The provider has provision across multiple CAH2 subjects but, due to scope and data thresholds, has only a single CAH2 subject that is assessed.	The provider- and subject-level metrics differ. This is because the provider-level metrics include all students, including from subjects that are not assessed.

95. Although the criteria are tailored at provider and subject levels, it is envisaged that a single-subject provider – where the entirety of provision is within a single subject – would always receive the same rating at provider and subject levels. We have considered two options for assessing single subject providers in the pilot: either the Main Panel could conduct the assessment using provider-level evidence, and the same rating would be awarded to the subject; or the subject panel could conduct the assessment using subject-level evidence and the same rating would be awarded to the provider.
96. Following initial feedback we will pilot the latter approach. Single-subject providers will write a maximum two-page provider summary statement and a maximum five-page subject submission, which are assessed by the relevant subject panel against the subject-level criteria and ratings descriptors. The subject rating, as determined by the subject panel, will also be awarded to the provider.
97. A provider-level perspective will inform the assessment, through the involvement of the chair or the deputy chair of the relevant subject panel (who are also members of the Main Panel).
98. ‘Multi-subject’ providers with only one assessed subject will be treated the same as providers with multiple subjects being assessed. They should write a maximum 15-page provider-level submission, a maximum two-page provider summary statement and a maximum five-page subject level submission for the assessable subject, and these will be assessed by the Main Panel and relevant subject panel. Due to the differences in both the data and the submissions, it is possible that different ratings could be awarded to the provider and its assessed subject.

Evidence used in the assessments

99. The TEF panels will make holistic judgements about a provider or subject's performance, based on the TEF metrics as well as additional evidence submitted by the provider, and in light of contextual information.
100. This section:
- provides a brief overview of the data and metrics that will be supplied by the OfS for use as evidence in the pilot assessments
 - provides guidance to participating providers on the purpose, scope and content of submissions they will need to prepare
 - provides guidance on involving students in submissions.
101. Full details of how the data and metrics are defined and constructed for the pilot are given in the 'Guide to subject-level pilot data'.¹⁹
102. Guidance on how the metrics and submissions should be interpreted and used by panel members to inform the assessments is given in paragraphs 231-284.
103. Figure 5 summarises the information used in subject- and provider-level assessment.

¹⁹ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

Figure 5: Information used in subject-level and provider-level assessment

		Provider level	Subject level
Evidence	Contextual information	<ul style="list-style-type: none"> Provider-level contextual data: <ul style="list-style-type: none"> profile of all students at the provider number of students taught in each subject Maps of provider-level geographic context 	<ul style="list-style-type: none"> Subject-level contextual data: <ul style="list-style-type: none"> profile of students studying the subject details of courses covered Maps of subject-level geographic context
	Metrics	Provider-level metrics: <ul style="list-style-type: none"> Teaching on my course (NSS) Assessment and feedback (NSS) <ul style="list-style-type: none"> Student voice (NSS) Academic support (NSS) Learning resources (NSS) Continuation (HESA/ILR) Highly-skilled employment or higher study (DLHE) <ul style="list-style-type: none"> Sustained employment or further study (LEO) Above median earnings threshold or higher study (LEO) 	
	Supplementary data	Where applicable: <ul style="list-style-type: none"> Differential degree attainment data and grade inflation data Additional data on part-time provision 	

Contextual information

104. Panel members will be supplied with contextual information which will allow them to understand the nature, range and operating context of the provision which they are assessing (including size, location, and student population), as well as aiding the interpretation of related metrics. Contextual data is used to support interpretation of performance but does not itself form the basis of any judgement. Providers will also have access to the contextual data to help draft their written submissions.
105. In the pilot we are testing an expanded set of contextual data, especially to make clear to subject panels what range of courses are covered within a subject and the extent to which they are interdisciplinary; and at both subject and provider level to improve contextual data relating to employment outcomes.
106. The contextual information will be tailored for subject- and provider-level assessment:
- a. Contextual information at provider level includes a breakdown of the number of students in each subject.
 - b. Contextual information at subject level includes information about the courses covered by the subject metrics and the extent to which they span more than one CAH2 subject.

107. Figure 6 sets out what contextual information will be provided to support provider-level and subject-level assessments.
108. We will also test whether provider-level contextual information is useful to inform subject-level assessment, after the initial assessment based purely on subject-level data.

TEF metrics

109. Metrics based on nationally collected data play a key role in informing TEF assessments. A number of datasets are used to produce a basket of metrics, which will be used at provider level and subject level. The metrics provide proxies of performance against the TEF criteria, as set out in Table 2.
110. Annex B provides full metrics descriptions, and the 'Guide to subject-level pilot data'²⁰ gives more detailed information about TEF metrics. It is recommended that the reader refers to the 'Guide to subject-level pilot data' throughout this guidance.
111. The metrics are presented as **core metrics**, aggregated to cover all students that are captured by each metric, and **split metrics**, which break down the core metrics into different groups of students.

What are the core metrics?

112. In the pilot we will test an expanded basket of nine core metrics, comprising three metric types:
 - a. Metrics based on the NSS:
 - i. The proportion of students who report being satisfied with 'teaching on my course' (as measured by responses to NSS Questions 1 to 4²¹).
 - ii. The proportion of students who are satisfied with the assessment and feedback they receive (NSS Questions 8 to 11).
 - iii. The proportion of students who are satisfied with the way their provider collects and responds to the student voice (NSS Questions 23 to 25).
 - iv. The proportion of students who are satisfied with the academic support they receive (NSS Questions 12 to 14).
 - v. The proportion of students who are satisfied with how far learning resources have supported their learning (NSS Questions 18 to 20).
 - b. Continuation metric:

²⁰ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

²¹ These are the question references as in the National Student Survey 2017 and the National Student Survey 2018. However, the question numbers differed in previous surveys, the data from which has also been used to inform TEF data. For more information, see Annex B of the 'Guide to subject-level pilot data' (See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/).

- i. The proportion of students who continue after their first year of study, as measured through analysis of administrative student data collections.
- c. Employment outcome metrics:
- i. The proportion of students in highly skilled employment or further study at a higher level six months after graduation, as measured by responses to the DLHE survey.
 - ii. The proportion of graduates in sustained employment or further study three years after graduation, as measured through analysis of the LEO dataset.
 - iii. The proportion of graduates earning above the median earnings threshold for 25-29-year-olds or further study at a higher level three years after graduation, as measured through analysis of the LEO dataset.
113. 'Student voice' and 'learning resources' will be included as two new core metrics to be tested in the second pilot. This decision has been made in response to feedback from students in the first subject pilot, to strengthen the metrics in the areas that matter most to students, and also reflects the fact that multiple years of data for these questions are now available for the first time, following changes to the NSS survey in 2017.
114. The two LEO-based metrics, included in previous TEF exercises as 'supplementary' metrics, will be tested as core metrics alongside the DLHE-based metric that focuses on highly skilled employment, as set out in the Government's consultation response²².
115. The LEO dataset has been enhanced since the production of metrics for the first subject pilot, such that three years of available data may now be drawn upon, and information on self-assessed employment and earnings is available and consistently defined. Testing LEO metrics as core metrics in this pilot also anticipates the discontinuation of the DLHE survey beyond 2018, and would provide continuity in the employment outcome metrics until new graduate outcomes data becomes fully available.
116. The DLHE-based core metric used in previous TEF exercises, measuring the proportion of students in employment or further study, has been removed. This ensures employment outcomes are not duplicated or over-weighted, and focuses on proxies for graduate-level employment outcomes.

²² www.gov.uk/government/consultations/teaching-excellence-and-student-outcomes-framework-subject-level.

Table 7: How TEF metrics map to aspects of quality

Evidence (TEF metrics)	Teaching quality (TQ)	Learning environment (LE)	Student outcomes and learning gain (SO)
Core metrics and split metrics (source)	Teaching on my course (NSS)	Academic support (NSS)	Highly skilled employment or higher study (DLHE)
	Assessment and feedback (NSS)	Learning resources (NSS)	Sustained employment or further study (LEO)
	Student voice (NSS)	Continuation (HESA/Individualised Learner Record)	Above median earnings threshold or higher study (LEO)
Supplementary data (at provider level only)	Grade inflation		Differential degree attainment

117. The same basket of core metrics will be used for both provider-level and subject-level assessment, aside from the supplementary data which will only be considered at provider-level.

Split metrics

118. The core metrics aggregate the data for all students covered by that metric, over a three-year period. To indicate performance relating to different groups of students or areas of provision, the metrics are also split by:

- each year of data (up to three years)
- student characteristics:
 - level of study
 - age
 - sex
 - disadvantage based on Participation of Local Areas (POLAR) (a measure of participation in higher education)
 - disadvantage based on national indices of multiple deprivation
 - Welsh-medium teaching
 - disability
 - ethnicity
 - domicile.

119. The provider-level metrics are also split by subject.
120. For more information about the splits, see the 'Guide to subject-level pilot data'¹²³ paragraphs 97-104.

Benchmarking and flags

121. Each core and split metric takes into account outcomes for similar students in the rest of the UK higher education sector to determine each provider or subject's **expected performance**, or 'benchmark' for their particular students.
122. This enables more meaningful interpretation of a provider's actual performance, and ensures that factors which may have an effect on student outcomes which are outside of the control of a provider are taken into consideration for assessment purposes.
123. The benchmark for a provider also takes account of the mix of subjects studied by its students. For most metrics, the benchmark for a subject is based only on students studying that subject across the sector. For the continuation metric it is based only on students studying in a group of subject areas²⁴ across the sector.
124. For each metric, a provider or subject's actual performance is compared to its benchmark. Two tests are applied to the difference between the two values to highlight where performance is above or below benchmark:
 - a. Materiality – whether the difference between a provider's actual performance and its benchmark is large enough to be considered 'material'. This is defined as a difference of 2 percentage points or more. The materiality test is not applied when the benchmark is above 97 per cent.
 - b. Significance – whether the difference between the two values is large enough that it is unlikely to be due to random chance. We use a formal statistical test that provides at least 95 per cent confidence that the difference is not due to random chance.
125. Where the actual performance of a provider or subject is both significantly and materially above or below its benchmark for that metric, a positive or negative 'flag' will clearly show this. A provider or subject's flags in its core metrics are used to calculate the starting point for each assessment.

Very high or low absolute values

126. In addition to the 'flags', which signal performance relative to the benchmark, very high or very low **absolute performance** will also be clearly marked. To define very high or low absolute values for we identify the values for the top and bottom 10 per cent of larger providers²⁵ in the sector, for each metric.
127. An asterisk (*) or an exclamation mark (!) will indicate that absolute performance is very high or very low. These markers are not benchmarked and do not take into account

²³ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

²⁴ The continuation metric benchmarks subject of study using nine groupings of the CAH2 subjects, instead of the 34 individual CAH2 subjects which are used in benchmarking all other metrics.

²⁵ Small providers with no more than 100 students in the population covered by the metric are not included in the calculations to identify the top and bottom 10 per cent.

differences between subjects. For the pilot, a statistical test is applied that identifies if there is at least 95 per cent confidence that performance is not marked as being in the top or bottom 10 per cent due to random chance.

128. Full details of how benchmarks, flags and absolute values are calculated and presented are set out in the 'Guide to subject-level pilot data'²⁶. Details of how they are used in assessment are set out in paragraphs 235-278 of this guide.

Other key features

129. The metrics for full-time and part-time students will be displayed separately. One of these will be identified as the majority mode of delivery for each provider and subject.
130. The metrics use three sequential years of aggregated data, taken from the most recent three years of available data. Where three years of data are not available, this will be clearly presented.
131. For each metric to be 'reportable' it must include at least 10 students, and meet the other reportability criteria, including response rate thresholds for the metrics drawn from survey data. For more information see paragraphs 65-70 of the 'Guide to subject-level pilot data'.

Supplementary degree attainment data (at provider level only)

132. In the pilot we will test an expanded set of supplementary data relating to degree classifications. This will be used at provider level only, and applies only to providers that hold taught degree awarding powers. The data will include Level 6+ undergraduate degree awards made by the provider to the students it has taught. The data will comprise:
 - a. Data on **differential degree attainment**. This will show the number and proportion of degrees awarded as 1sts and 2:1s, six, four and two years ago, to each of the student groups shown in the splits. This data is at an early stage of development and we expect in future to align it with the OfS's approach to access and participation in higher education. The data will not be benchmarked, but to aid interpretation the panel will be provided with sector-wide data about differences in the degree classifications awarded to students from different backgrounds. We will explore how this data can be developed and used in submissions alongside the splits and evidence, in relation to assessing 'Positive outcomes for all'.
 - b. Data on **grade inflation**. This will show the number and proportion of degrees awarded as 1sts, 2:1s, other degree classifications and unclassified degree awards 10, six, four and two years ago. As in provider-level TEF, It will be used in combination with evidence in the submission as potential evidence in relation to assessing 'Rigour and stretch'. This data will also not be benchmarked, but to help contextualise it, we will include data on students' entry qualifications for the cohorts graduating six, four and two years ago, and we will provide the panel with sector wide data relating to grade inflation. The differential degree attainment data may also help contextualise changes in a provider's overall grade profile.

²⁶ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

133. Further definitional information about the supplementary data is provided in the Guide to subject-level pilot data at paragraphs 105-109.

Additional data about part-time provision (at provider-level only)

134. Providers with at least 35 per cent of their students studying part-time may, for the provider-level assessment only, submit an additional page of quantitative information which will be considered alongside the assessment of the metrics in step 1b of the assessment process (see paragraphs 264-265).
135. The purpose of the information is to supplement the metrics where limitations specifically affect part-time provision. In particular the data should relate to:
- student satisfaction, for part-time students who are not in scope of the NSS (that is, students on courses of 1 full-time equivalent or less, such as HNCs)
 - continuation, for part-time students who are studying flexibly
 - employment, for part-time students already in work or not seeking work after their studies.
136. Information must be quantitative, and may derive from the provider's own records or from external sources of information. Although the information need not follow a standardised format, where possible it should attempt to place the data in the context of national comparators.
137. Submission of this additional information about part-time provision is optional.

Access to the metrics

138. The provider- and subject-level metrics will be issued both to participating providers to help them to submit written evidence, and to panel members to support their TEF assessments.
139. We will release subject-level pilot data to all providers that are eligible for TEF Year Four to support the sector and its students to prepare for subject-level TEF. This pilot remains formative with regard to its role in developing the government's policy and design for subject-level TEF: as such it will be conducted in confidence, with no individual ratings generated from the pilot being published. The formative and developmental nature of this pilot means that the TEF metrics and contextual data defined for this purpose remain experimental and are intended as a tool for constructive and collaborative policy development. While subject-level TEF remains in development, the OfS consider that publication of the TEF metrics being piloted would be prejudicial to the effective development of government policy. Consequently, the OfS does not plan to publish this year's pilot TEF metrics and contextual data. The information will be shared with providers solely for the purposes of participating in this pilot or engaging with the government's development of subject-level TEF more generally.
140. On receipt of the pilot TEF metrics and contextual data, all providers will be instructed that the information must not be disclosed to any third parties, including staff or students who are not directly involved in the production of TEF submissions or the development of subject-level TEF policies and processes. A record of the individuals to whom access has

been granted must be made and maintained by the provider, to include the purpose of such access for the individual in question.

Submissions

141. The evidence used in TEF assessments comprises metrics as well as written submissions, which are essential for the panels to make holistic assessments against the broad set of TEF criteria.
142. Three types of written submission are required in the pilot:
- a provider submission (no more than 15 pages)
 - a subject submission (no more than five pages) for each subject being assessed
 - a single provider summary statement (no more than two pages) that will be made available to subject panels (see paragraphs 171-175 for more information).
143. A subject submission is required for every subject that is being assessed; that is, every subject that is in scope and has sufficient data for assessment.
144. In addition, some providers are eligible to submit an optional extra page of quantitative information relating to part-time students (see paragraphs 264-265).
145. Single-subject providers will submit a subject submission plus provider summary statement. Please refer to paragraphs 94-98 for more details.
146. Table 8 gives an overview of the differences between provider and subject-level submissions.

Table 8: Features of provider- and subject-level submissions

Provider submission	Subject submissions
<ul style="list-style-type: none"> • 15-page maximum • Addresses provider-level criteria • Engages with provider metrics (and supplementary degree attainment data where applicable). 	<ul style="list-style-type: none"> • Five-page maximum • Addresses subject-level criteria • Engages with subject metrics. <p>In addition providers should submit a provider summary statement of up to two pages maximum.</p>

Purpose and scope of provider and subject submissions

147. The purpose of submissions is to ensure that judgements are based on evidence against the TEF criteria that is broader than metrics information alone. Provider and subject submissions will be used by panel members alongside the TEF metrics to inform holistic judgements. Therefore, **provider and subject submissions are compulsory**.
148. Submissions should:
- provide further context
 - engage with the metrics

- provide additional evidence in relation to all three aspects of quality.
149. In this pilot we will test guidance that is more directive about the scope and content of submissions, so that they can be assessed in a comparable way. In particular, the guidance in paragraphs 159-170 sets out clearer expectations about the range of additional evidence that should be included, and how submissions should engage with the metrics. However, we do not provide a prescriptive template for submissions, recognising the diversity of providers and subjects and the many ways in which evidence of excellence can be presented²⁷.
 150. In this pilot we also set clear expectations that students should be given meaningful opportunities to be involved in creating both provider and subject submissions. The student declaration (see paragraphs 178-185) aims to provide assurance to the panel that this took place. Submissions should explain how students have been involved, and any student-authored sections should be clearly signalled.
 151. To enable the panels to make holistic best-fit judgements against the rating descriptors, submissions should at minimum cover **all three aspects of quality** (Teaching Quality, Learning Engagement and Student Outcomes), engaging with the metrics and providing additional evidence in relation to each aspect. Submissions do not need to address every criterion.
 152. Particularly at the subject level, where the page limit is more restricted, submissions should focus on those criteria that best help make a case for excellence, alongside the available metrics.
 153. Provider- and subject-level submissions are expected to cover provision for all students in scope of the assessment. This includes both full-time and part-time students, and international students studying in the UK or other groups who may not be captured in all the TEF metrics for technical reasons, but nevertheless remain in scope of the assessment.
 154. Subject-level submissions should cover all programmes and modules that make up the metrics reported for the subject, even where these cross internal departmental or other organisational lines.
 155. Where providers wish to draw out the impact of their interdisciplinary approach separately from the TEF metrics, they will be able to do so through the relevant written submissions, for example in each of the individual subjects involved.
 156. Evidence included in the written submissions should relate to provision and impact during the last three years, except where explaining metrics related to earlier years. Any plans for future initiatives are out of scope.
 157. Where there are franchise arrangements, the quality of provision will be assessed at the provider that delivers the teaching. This may or may not be the provider that awards the qualification or registers the student. Submissions should cover all undergraduate teaching

²⁷ Guidance on the format and layout of submission documents is at Annex B, 'Additional information for participating providers and TEF student representatives'. A broad template that is correctly formatted will be available, but providers may if they wish adapt the headings.

delivered by the submitting provider, although the metrics will associate students with the provider where the student spends the majority of their first two years of study.

158. Provider and subject submissions made in this pilot should be written as if they would be published, as they would be in a real exercise.

Additional evidence

159. The final rating awarded is based on the panel's holistic judgement of the evidence in the metrics and the submission. Providers should note that the rating given by the initial hypothesis can move down as well as up when the totality of evidence is considered against the rating descriptors and a best-fit judgement against them is made. For example:
 - a. Where the initial hypothesis based on the metrics is Gold (or Silver), the submission will need to include additional evidence in the submission that supports that hypothesis, to merit a best-fit judgement of Gold (or Silver) based on the totality of evidence.
 - b. Where the initial hypothesis 'defaults' to Silver largely due to limitations in the data (such as non-reportable metrics or a lack of flags due to small cohorts) the submission will need to provide substantive additional evidence that aligns with the Silver (or Gold) rating descriptor, to merit a best-fit judgement of Silver (or Gold).
160. Submissions can include both quantitative and qualitative evidence, drawn from internal or external sources. Evidence should be directly relevant to the TEF assessment criteria, and should focus on demonstrating excellence rather than meeting baseline or threshold expectations.
161. Panels are particularly looking for evidence of the impact that policies and practices have had on student learning and outcomes, rather than descriptions of policies and practices alone. Submissions should connect policies, practices and the impact they have had, referring to the TEF metrics or other evidence of impact supplied within the submission.
162. The balance between policy and practice may be different at the provider and the subject level: for instance, a subject submission may focus on how the subject implements the provider-level policies and helps students engage with the resources set out in the provider summary statement.
163. Table 9 gives examples of the types of evidence that may be included in submissions at provider and subject levels. These examples are **not prescriptive or exhaustive**. A submission may include any type of evidence that is within scope of the TEF.

Table 9: Examples of types of evidence in submissions

Provider level	Subject level
Types of evidence applicable to any of the aspects	
<ul style="list-style-type: none"> • Explanations of how provider-wide policies inform practice, and evaluation of their impact. • Evidence used for internal monitoring and evaluation of institutional strategy or provider-wide initiatives, where relevant to the TEF criteria. • Internal data based on surveys, monitoring, learning analytics, etc., with explanation of how the data is representative across the provider. Where appropriate, trend data should be included to indicate the impact of initiatives. • Commentary, analysis or collated feedback from students. Explanation of how this is representative of students across the provider should be included. • Excerpts from Quality Assurance Agency for Higher Education reviews, where they indicate performance above the baseline. 	<ul style="list-style-type: none"> • Explanations of pedagogic approaches, enhancement activities, or innovations in teaching and learning, and evaluation of their effectiveness. • Evidence used for internal programme, departmental or subject-related reviews, where relevant to the TEF criteria. • Internal data based on surveys, monitoring, learning analytics, etc., with explanation of how the data is representative across the subject. Where appropriate trends data should be included to indicate the impact of initiatives. • Commentary, analysis or collated feedback from students. Explanation of how this is representative of students across the subject should be included. • Accreditation of courses by PSRBs or excerpts from PSRB reports, where they indicate performance above threshold requirements.
Teaching quality (TQ)	
<ul style="list-style-type: none"> • Monitoring and analysis of students' engagement with their studies, for example through the UK Engagement Survey or learning analytics data. • Explanation of how provider-wide policies or schemes for recognising and rewarding teaching excellence inform practice, and evaluation of their effectiveness, for example trend data on progression and promotion of staff based on teaching performance. • Data relating to the qualifications, experience, and external recognition of staff who teach. • Evaluation of the effectiveness of external examiner and grading systems, or other policies, in maintaining standards. • Analysis of trends in degree attainment over time, and how these relate to institutional policies or other factors. • Explanation of how provider-wide policies for involving students as partners inform practice, and evidence of their implementation across the provider. 	<ul style="list-style-type: none"> • Monitoring and analysis of students' engagement with their studies, for example through the UK Engagement Survey) or learning analytics data. • Data on the amount and types of contact time and student engagement with independent learning. • Explanation of how excellent teaching is valued and supported in the subject and evaluative data, for example on progression and promotion of staff, external recognition or awards, or the effectiveness of teaching observation schemes. • Excerpts from external examiner reports representing provision across the subject. • Analysis of internal student feedback, for example module evaluations, or feedback gathered from course representatives. • Impact and effectiveness of feedback initiatives aimed at supporting students' development, progression and achievement.

<ul style="list-style-type: none"> • Examples of provider-wide initiatives driven by student input. • Representative student commentary on students' involvement in and influence on the development of provision or institutional strategy. 	<ul style="list-style-type: none"> • Evidence of how student feedback is collected and acted on. • Explanation of how students are involved as partners, for example in course or curriculum development, and evidence of their influence.
Learning environment (LE)	
<ul style="list-style-type: none"> • Data about investment in teaching and learning resources and evidence of the impact of such investments. • Data on students' use of learning resources and relevant support services, and evaluation of their impact on student learning. • Explanation of the provider's strategy for engaging employers and practitioners in provision, with data on the extent to which this is implemented across the provider. • Information about the research context of the provider and evidence of how it is used to enhance learning and teaching. • Data on completion and attainment of all groups of students, and evaluation of how provider-wide strategies or initiatives have impacted on these. • Analysis of the characteristics of the providers' students and their range of learning needs, and explanation of how these are supported. • For relevant providers, evidence of Welsh medium provision contributing to students' academic experiences. 	<ul style="list-style-type: none"> • Data about investment in subject-specific teaching and learning resources and evidence of the impact of such investments. • Data on students' engagement with learning resources and independent study. • Information about the range of employers and practitioners engaged in course design and delivery, and representative feedback from employers and practitioners. • Data about the extent of student involvement in or exposure to the latest developments in research, scholarship and professional practice. • Data on completion and attainment of all groups of students at subject-level, and evaluation of how practices in the subject have impacted on these. • Information about the range of students' learning needs, and explanation of how these are supported.
Student outcomes and learning gain (SO)	
<ul style="list-style-type: none"> • Internal measures or analysis of learning gain. • Analysis of graduates' destinations and longer-term employment outcomes and progression. • Explanation of provider-wide strategies, initiatives or services aimed at enhancing employability or progression to employment, and evaluation of their impact. • Data or analysis of career enhancement for mature students. • Data about graduate start-ups or student involvement in enterprise and entrepreneurship. • Analysis of gaps in attainment of disadvantaged groups; explanation of policies and initiatives aimed at closing 	<ul style="list-style-type: none"> • Internal measures or analysis of learning gain within the context of the subject. • Analysis of graduates' destinations and longer-term employment outcomes and progression. • Explanation of how employability is embedded within the curriculum and wider student experience. • Representative feedback from graduates about how their education has enhanced their careers and personal lives. • Representative feedback from employers about graduates' employability and skills. • Data about graduate start-ups or student involvement in enterprise and entrepreneurship.

gaps; and evaluation of their effectiveness.	<ul style="list-style-type: none"> • Analysis of gaps in attainment of disadvantaged groups; explanation of practices aimed at closing gaps; and evaluation of their effectiveness.
--	--

164. The quality and relevance of the information that the provider submits, including evidence from student representatives, are subject to the scrutiny of the panel. In the pilot we will include a process for panel members to request, in exceptional circumstances, clarification or verification of particular information contained in a submission, where the response might materially affect the overall judgement. Through the pilot we will evaluate whether such a process will be necessary to retain in a full subject-level TEF exercise.

Engaging with metrics

165. As well as providing additional evidence against the criteria, submissions should engage with the metrics across all three aspects of quality, and where possible address limitations in the data.

166. It is particularly important for submissions to:

- a. Explain how performance indicated by the metrics relates to the providers' policies and practices, alongside other evidence of their impact.
- b. Address negative flags in the core or split metrics, or low absolute values. Providers may wish to:
 - i. Explain what is causing the negative flags.
 - ii. Set out any mitigating factors that explain the metrics.
 - iii. Set out any internal data that supplement, explain or contradict the metrics.
 - iv. Explain what actions are under way to improve performance, providing evidence of their impact where possible.
- c. Fill gaps for non-reportable metrics. Providers should supply internal data where possible to fill in gaps, especially in the majority mode of delivery. For example, a subject with non-reportable NSS metrics could supply internal evidence of student satisfaction such as student end-of-module evaluations.
- d. Supplement metrics in the majority mode of delivery that are based on very small student numbers, and have low levels of statistical significance. In these cases the panel will need to rely heavily on evidence within the submission.
- e. Include the following where NSS data has been materially affected by a boycott:
 - i. If the boycott has caused the NSS metrics to be non-reportable, any possible alternative sources of evidence against the relevant criteria. This may be drawn from internal student feedback, or other forms of evidence related to these criteria.

- ii. If the provider considers that its NSS metrics have been materially affected by the boycott, the submission should include evidence of this. The submission should also include, if possible, alternative sources of evidence against the relevant criteria.
 - f. Fill gaps, where possible, for any significant student populations excluded from the metrics (for example, international students who are not included in the employment metrics).
167. We will work with providers and panel members through the pilot to explore how submissions can best mitigate limitations in the data, especially where there are non-reportable metrics, very small cohorts, or NSS data affected by boycotts.
168. Providers may also refer to contextual information in their written submissions, and may wish to draw out the implications of particular contextual information, for example to explain how it may affect their metrics performance or other aspects of their provision.
169. Providers may wish to draw panel members' attention to the year splits in the metrics, as potential evidence that policies and practices cited in the submission are having an impact.
170. In addition, at provider level only, the submission should engage with the supplementary degree attainment data (where applicable). The submission should explain any significant changes in the overall profile of degree classifications awarded, and describe how attainment gaps are identified, policies aimed at closing gaps and their impact.

Provider summary statement

171. Feedback from the first pilot indicated that many providers wished to use space in their subject submissions to set out the provider's context to effectively explain their subject-level provision, which reduced the amount of space available to provide evidence for excellence at subject level. Subject panels often found this helpful in their assessments.
172. Therefore in this pilot we will test the inclusion of a separate provider summary statement; a single statement from each provider that will be made available to the subject panels to inform their subject assessments. This is intended to enable providers to focus their five-page subject submissions more effectively on the case for excellence at subject level, and set this in the institutional context. We will test how this enables the subject panels to assess subjects in the context of the provider, its policies and institutional facilities that help support teaching and learning within the subject.
173. The statement may include qualitative and quantitative information about:
- the mission and relevant aspects of the provider's external environment and operating context
 - the provider's internal structure, and how responsibilities for oversight and delivery are centralised or delegated between provider and subject levels
 - Institutional policies and facilities that help support teaching and learning within the subject.

174. The provider summary statement should help panels situate the subject within the institutional context. It may also provide evidence in relation to the quality or effectiveness of institutional policies or facilities. The subject submission should then explain how these policies are implemented or facilities are made use of within the subject, and demonstrate the impact on the students' experiences and outcomes.
175. The provider summary statement should summarise, in no more than two pages, relevant content from the lengthier provider-level submission. The summary statement will not be made available to the Main Panel to inform the provider-level assessment, so must not include any additional evidence of excellence that is not contained in the provider-level submission.

Student engagement with submissions

176. Providers are expected to provide opportunities for and seek to secure meaningful student engagement with the provider- and subject-level submissions.
177. Pilot participants are expected to offer the following types of opportunities for student engagement with the provider and subject-level submissions. Providers are expected to work with their nominated TEF student representative (see paragraph 179) to agree on approaches which may work within their particular context, and may consider, for example:
 - a. An authentic partnership approach to considering evidence and performance against the relevant criteria.
 - b. Involvement of student representatives in TEF-related steering and working groups throughout the pilot, with a clear mandate to help critique and shape, rather than simply validate the provider response.
 - c. Deliberate engagement with the wider study body on TEF. Often this will be through, but beyond student representative bodies (such as students' unions or equivalent, and course representatives, or equivalent).
 - d. Sharing of TEF metrics information in full at the earliest opportunity, to enable their engagement with the evidence and contribution to submissions.
 - e. Support for student understanding of the TEF and metrics information by providing guidance, offering workshops and organising staff support.
 - f. Invitations to students, preferably through their representatives, to contribute their own comments and evidence to inform or be included in the submissions.

TEF student representative and the student declaration

178. The pilot will test ways in which students can best be supported by providers and the OfS to engage with TEF and contribute to the evidence that will be considered as part of subject-level TEF assessments.
179. The provider will be asked to nominate a TEF student representative (for example, a student union president or relevant sabbatical officer or equivalent), and a TEF student alternative contact (for example, a staff member from the student union or equivalent) at the

outset of the subject-level pilot. Providers are expected to work with their existing student representative bodies or equivalent to identify these contacts.

180. The TEF student representative should have opportunities to:
- be involved in the provider submission, should they wish to be
 - access to the pilot TEF metrics, via the provider, to support their involvement in submissions
 - help co-ordinate any student involvement in provider-level and subject-level submissions
 - attend pilot briefing and feedback events run by the OfS
 - confirm to the OfS whether students have had appropriate opportunities for involvement in the pilot submissions
 - provide evaluative feedback to the OfS.
181. The OfS will support TEF student representatives, including through events to provide training and briefings (this will include a ‘train the trainer’ event to support them in cascading information to subject-level student representatives).
182. We wish to encourage active student engagement in the pilot, and expect participating providers to ensure there are meaningful opportunities for involvement. To test ways of ensuring that such opportunities are provided, we will ask the TEF student representative to provide a declaration of this as part of the submission process. The declaration will be made available to the pilot Main Panel.
183. If the representative student union or equivalent organisation is unwilling or unable to participate in the pilot, the provider should still seek to nominate the appropriate individual as the TEF student representative. We will ask them to note that they are unwilling or unable to participate in the student declaration, and will seek their feedback on barriers to participation.
184. The provider may still seek other ways to involve students with the process if appropriate, and may wish to explain how they have done so in their submissions. The provider will not be disadvantaged in the assessment if it has made appropriate opportunities available but its student representatives declare that they have chosen not to be involved.
185. The OfS will invite TEF student representatives to provide feedback on the pilot, to inform our future approach to, and support for, student involvement in subject-level TEF. In gathering feedback we will bring together student panel members and TEF student representatives from participating providers to share views across all aspects of the process, and identify ways in which subject-level TEF can be developed further to focus on what matters most to students.

The assessment process

186. The assessments are carried out by independent panels of academics, students, and subject and other experts. They follow a three-step method of making holistic assessments

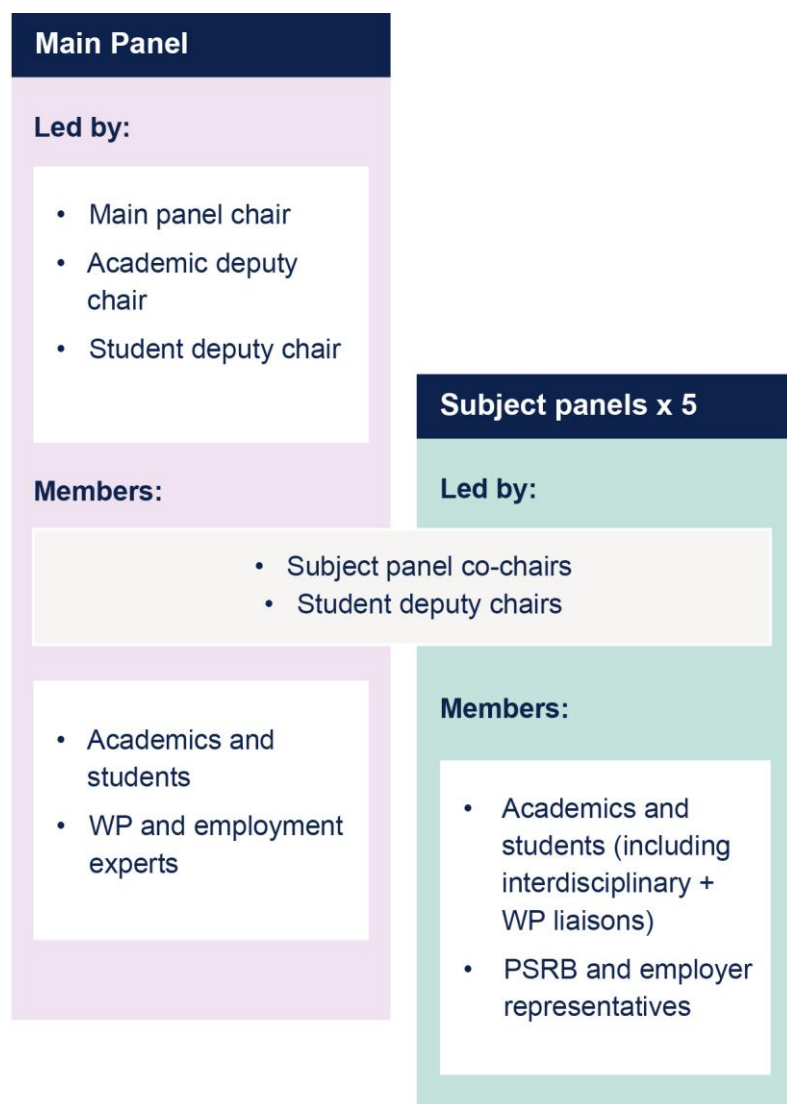
against the criteria, taking into account all the evidence available in the metrics and submissions.

187. A key challenge for full subject-level TEF will be to carry out robust assessments in a consistent manner, on a much larger scale. In this pilot we are explicitly testing panel processes and a panel structure that could be deployed to accommodate a significantly larger volume of assessments in full subject-level TEF, including:
- a revised configuration of subject panels, which can be adapted to accommodate larger volumes
 - a scalable process for carrying out the assessments
 - mechanisms for ensuring consistency across subject panels.

The pilot panels

188. The pilot Main Panel will be collectively responsible for deciding the provider-level ratings.
189. Each pilot subject panel will be collectively responsible for deciding the subject-level ratings for all subjects within its remit. Please see Table 5 for full information about the subjects that will be considered by each panel.
190. The Main Panel will provide oversight, and we will test mechanisms for the Main Panel to ensure consistent standards of TEF assessment across the subject panels
191. The pilot Main Panel will comprise: a main panel chair, a deputy academic chair, a deputy student chair, the co-chairs and student deputies from the subject panels, additional academic members and additional student members, and WP and employment experts.
192. The pilot Main Panel's key responsibilities are:
- producing provider-level ratings and statements of findings for all participating providers
 - providing oversight to ensure consistent standards of assessment by the subject panels, and support the assessment of WP issues and interdisciplinary provision
 - providing feedback on the pilot process from across the main and subject panels, and providing advice and recommendations on future subject-level TEF design and delivery.
193. Full subject-level TEF will require subject panels to make thousands of subject-level assessments. The configuration of pilot subject panels and the assessment processes are designed to test approaches that can be scaled up to a full subject-level exercise. At this stage we anticipate that full subject-level TEF will require approximately 10 subject panels.
194. To test scalable processes with realistic panel workloads, we have five 'paired' subject panels in the pilot. Each of the five 'paired' pilot subject panels comprises two main clusters of subjects, and could operate as two separate subject panels in future.

Figure 6: Membership of pilot panels



195. Each pilot subject panel will comprise two co-chairs and two student deputy chairs (to cover each of the 'paired' subject clusters), approximately 12 academic members and six student members, and additional employer or PSRB representatives. For full subject-level TEF we anticipate that each of these panels would split into two, and each of the 10 resulting subject panels would expand its membership as needed to manage the volume of assessments.

196. The pilot subject panels' key responsibilities are to:

- produce subject-level ratings and statements of findings for all of the subjects within their remit
- provide feedback from the subject panel's perspective, and provide advice and recommendations to the Main Panel on future subject-level TEF design and delivery.

197. On each subject panel we will identify members who will act as WP liaisons. These panel members will identify and highlight widening participation issues within and across panels, working with the Main Panel experts to develop advice and ensure a consistent approach to assessment as necessary.

198. On each subject panel we will also identify members who will act as interdisciplinary liaisons. They will be academic or student members of the subject panel, with additional responsibility for advising other panel members in their discussions of cases concerning interdisciplinary provision, raising the panel's awareness of interdisciplinary provision, and liaising with other subject panels as appropriate. They will also participate in additional activities including providing feedback on the process.
199. A full list of panel roles and responsibilities is included at Annex D.

How assessment is organised

200. In the pilot we will test an approach to organising the assessments that builds on practices that the provider-level TEF panel has found to be effective in previous TEF exercises, and adapt these to be scalable for full subject-level TEF. For both the provider-level and subject-level assessments this involves:

- training and calibration exercises for all panel members to prepare them for assessment
- allocation of cases
- the following stages of assessment:

Stage 1: detailed assessment of each case by three members of the relevant panel

Stage 2: discussion by a larger group of panel members, to form a recommendation to the whole panel

Stage 3: consideration of all recommendations and final decisions by the panel as a whole, with particular attention given to borderline or more challenging cases.

Training and calibration

201. Before commencing assessment, all panel members will receive training initially through online webcasts, an online training module and a face-to-face training event. Training will cover:

- the criteria and assessment framework
- the evidence used in assessment, including technical understanding of the metrics
- the three-step method for making judgements, taking full consideration of all the available evidence
- how the Main Panel and subject panels will interact.

202. Panels will also undergo a calibration exercise to fully prepare members for the assessment. A small sample of submissions will be read for calibration purposes to allow panel members to discuss the assessment process, clarify uncertainties and develop a common understanding of the TEF criteria and ratings descriptors, and how they should be used to underpin judgements.

Case allocation

203. Panel chairs and deputies will work together to allocate material to panel members for assessment. Where chairs or the deputies have a conflict of interest they will not participate in the allocation of material for assessment.
204. Each case will be initially allocated to a 'trio' of panel members comprising two academic panel members and one student panel member. Care will be taken to allocate each case to a trio with a mix of specific and broader expertise:
 - a. At provider level, at least one panel member will have expertise relating to the type of provider.
 - b. At subject level, at least one panel member will have expertise relating to the particular subject.
205. Members with broader expertise (that is, of different types of providers, or different subjects within the subject panel's remit) will support critical understanding and consistency in judgements, and provide externality.
206. Similarly each panel member should be allocated a caseload of assessments which represents a variety of cases, for example different provider types or a range of subjects, and also cases with different initial hypothesis starting points. This is intended to give them experience in carrying out assessments across the rating scale.
207. For each case, one individual will be identified as the lead reviewer. This reviewer will be responsible for leading discussions in further stages of the assessment and drafting the statements of findings.
208. Chairs and deputies may also identify a number of cases for additional consideration by specialists, for example where they anticipate it will be instructive to gain input from employers or PSRB representatives, or from panel members with expertise in WP or interdisciplinarity. The WP and employment experts on the pilot Main Panel and the employer and PSRB representatives on the subject panels will review samples that cut across the trios' caseloads, to draw out general issues relating to WP and employment considerations as well as advise on specific cases.

Conflicts of interest

209. Panel members are required to declare conflicts of interest they hold with any providers participating in the pilot. Panel members will not be allocated and will not take part in assessing or recommending or deciding the outcome of any provider with which they hold a conflict of interest.
210. For the purposes of the pilot, conflicts of interest should be declared where, within the last five years:
 - the individual worked for or studied at the provider
 - the individual was a board or council member of the provider
 - the individual held an honorary position at the provider

- the provider was an awarding body or delivering partner of the individual's institution, and the individual had personal responsibility for or involvement in these arrangements.

Stage 1: Individual assessment

211. Following the allocations and the calibration exercise, each case will be reviewed in detail by the trio.
212. Individual reviewers in the trio will review each allocated case in advance of meetings to independently arrive at a recommended TEF rating (individual ratings may be made against the five-point scale). They may also conclude that there is limited evidence and that reaching a TEF rating is not appropriate.

Stage 2: Recommendations

213. Panels will meet initially in groups of trios (typically nine panel members), to discuss the Stage 1 ratings for their caseloads and arrive at recommendations for the whole panel. The group of nine may recommend a rating of Gold, Silver, Bronze, or a borderline rating, or that there is insufficient evidence to make a best-fit judgement, or report that the group was unable to reach a consensus view. Where the recommendation is not for a Gold, Silver or Bronze rating, additional panel members may be asked to review the submission in detail to inform discussions in Stage 3.
214. During these discussions the lead reviewer will introduce discussion and will note the conclusions and recommendations of the group.

Stage 3: Decisions and outcomes

215. The whole panel will convene to consider the Stage 2 recommendations and agree final ratings. All cases will be considered, with particular attention paid to borderline cases and cases which were flagged as particularly challenging. In all cases, the panel will aim to reach an agreed rating of Gold, Silver or Bronze, though they may also conclude that there is insufficient evidence to make a best-fit judgement on the rating.
216. The lead reviewer will note the main points agreed in the panel meeting and will use this as a basis for drafting a statement of findings to explain the rationale for the rating.
217. In Stage 3 of the provider-level assessments, the Main Panel will also consider the final set of subject-level ratings for each provider. We anticipate in the pilot this may affect the outcome either where the rating based on provider-level information is very close to a borderline, or where the profile of subject ratings appears to contradict the provider rating. In the latter case, the Main Panel may wish to reconsider the provider rating. We will test with the Main Panel whether this approach helps achieve coherence between the subject and provider-level assessments.

Clarification and verification

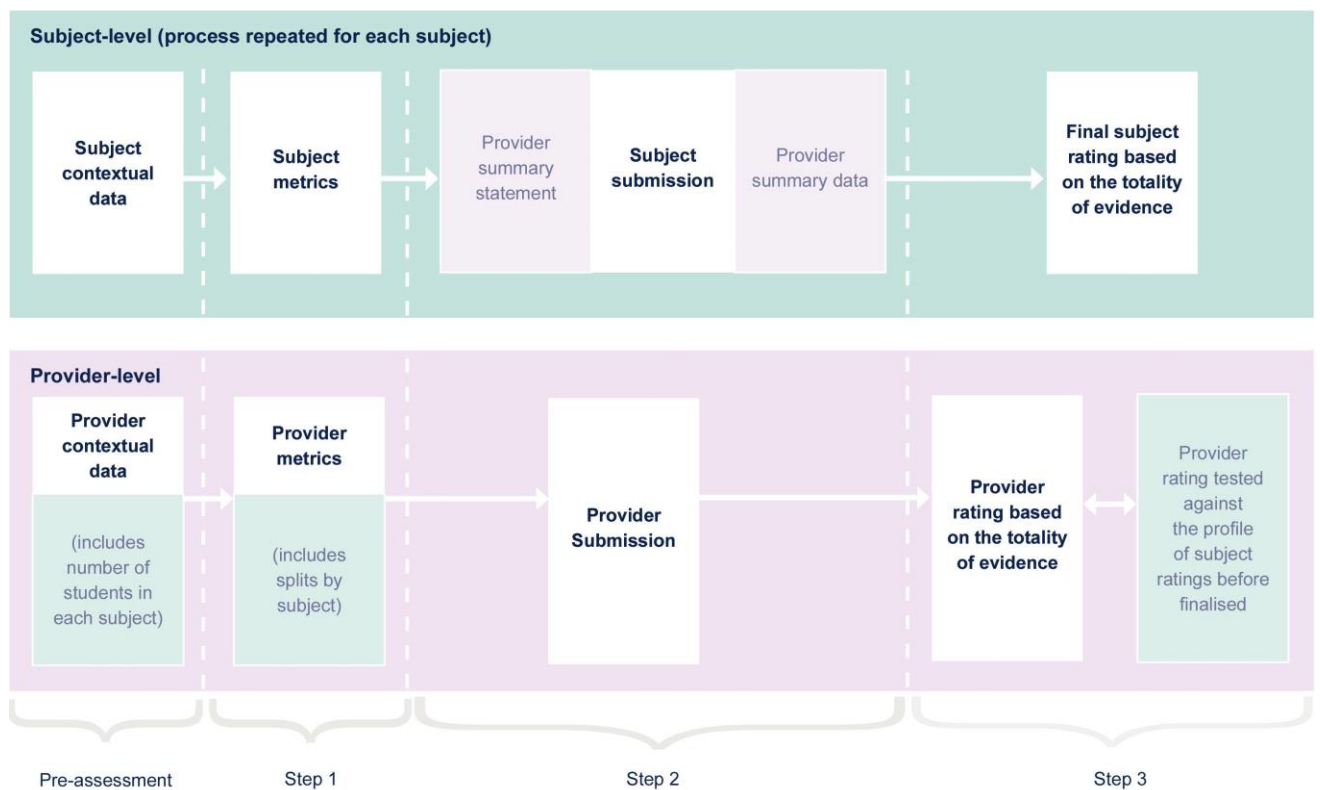
218. Panel members may, by exception, request clarification or verification of critical information already contained in a provider or subject submission. Panel members should raise queries only where the response could potentially make a material difference to the overall outcome.

219. Queries should be raised at the earliest possible opportunity via the TEF team, once the submission has been reviewed. The decision to request clarification or verification from the provider will be made at the discretion of the head of TEF or the TEF subject pilot manager. Once the TEF team receives the response it may redact content to remove information that is out of scope (for example, if the response seeks to expand on or add new evidence to a submission) before passing the response back to the requesting panel members.
220. The clarification and verification process will be kept under review throughout the pilot, with a view to establishing whether it is necessary to include this process in a future subject-level TEF exercise.

Relationship between subject- and provider-level assessment

221. We will test how information about the provider should be considered as part of the subject-level assessment; and how information about subjects should inform the provider-level assessment. This is summarised in Figure 7, and further detail is at paragraphs 286-288 and 293:

Figure 7: Information used in subject- and provider-level assessment

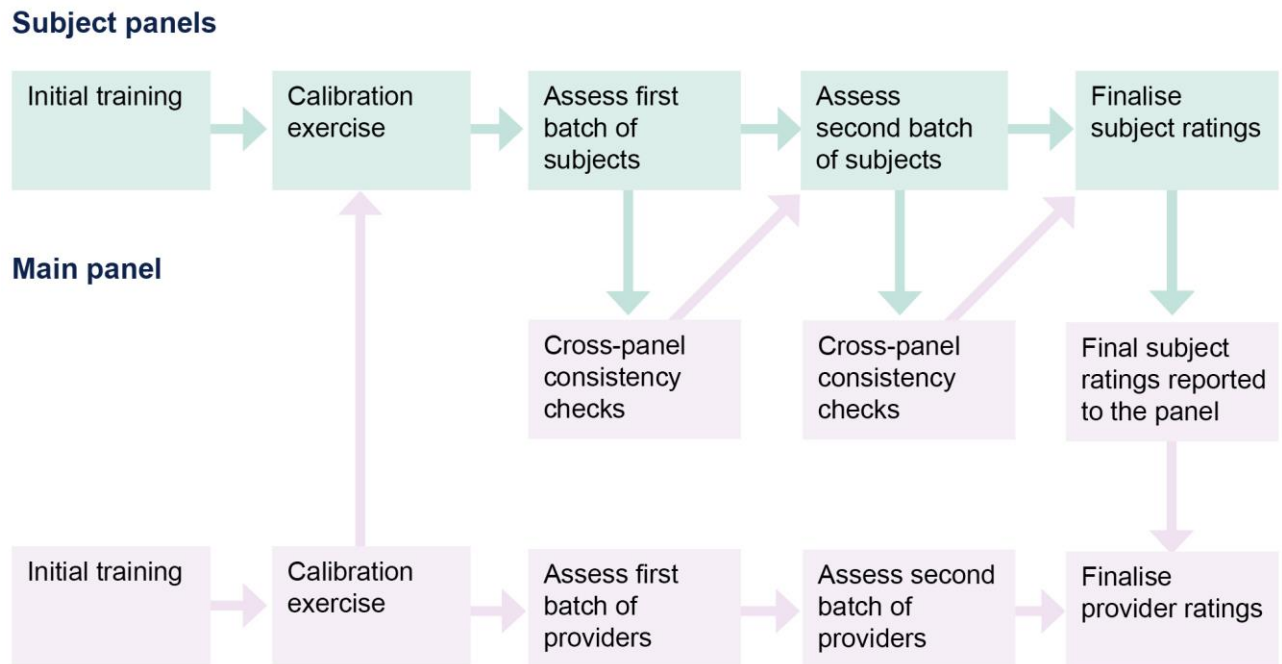


222. We will also test how the Main Panel and subject panels should interact, including the following, as set out in Figure 8:
- Overlap in membership.** Subject panel co-chairs and deputies will be members of the Main Panel, providing a link between the two.
 - Oversight by the Main Panel.** The Main Panel will check for consistency in the standards of assessment across the subject panels, and provide guidance to the subject panels. The subject panels will remain responsible for determining the

ratings at subject level, but will take account of advice and guidance from the Main Panel on the applications of consistent standards of assessment.

- c. **Specialist expertise.** The Main Panel includes specialist WP and employment experts, and the subject panels include members identified as WP and interdisciplinary liaisons. We will test how best to co-ordinate and deploy relevant expertise across the panels.

Figure 8: Interaction between Main Panel and subject panels



Moderation and consistency

- 223. The pilot will test processes for moderating judgements within each panel, and for ensuring consistent standards of assessment across the subject panels.

In-panel moderation

- 224. Subject panels and the Main Panel will carry out initial calibration exercises to develop a common understanding among its members of the assessment standards.
- 225. As the assessment progresses, each panel will monitor the profile of ratings across individuals, subjects and provider types, and will pay particular attention to the ratings for borderline or more complex cases. Panels may revisit ratings from earlier meetings to do this, with all ratings finalised only at the final meeting.

Cross-panel consistency

- 226. The Main Panel will test means of ensuring consistency of assessment standards across the subject panels.
- 227. Provider-level assessment will be conducted primarily by Main Panel members who are not subject panel chairs and deputies; and the consistency checks across panels will be carried out primarily by the subject panel chairs and deputies. There will be overlap, however, with

the whole panel taking collective responsibility for the provider ratings and for consistency checks.

228. Checks for consistency will include:

- a. **Progress reports from subject panels at each Main Panel meeting**, including opportunities to raise and resolve issues that caused uncertainty in the assessments.
- b. **Comparison of rating profiles across the subject panels provided to the Main Panel** after each round of subject panel meetings, and how they relate to metrics profiles. Where divergences are identified these will be discussed and may be investigated further or resolved through providing advice to the subject panels.
- c. **Observation at subject panel meetings by some Main Panel members**, including the chair and deputies as well as WP and employer experts. Their observations will be fed into discussions of the Main Panel.

Method of assessment

229. At each stage in the assessment process outlined in paragraphs 211-217, panel members will follow a three-step method of reviewing the available evidence and arriving at a rating. The same method will be used for provider- and subject-level assessment, although there are some differences in the information used at provider and subject level in particular steps, as detailed in Table 10:

- Pre-assessment: Consideration of contextual information.
- Step 1: Review the metrics to form an initial hypothesis. This starts with a formulaic method based on core metric flags (step 1a), and then involves reviewing all the metrics data to make a judgement about the initial hypothesis (step 1b).
- Step 2: Review of the submission, including consideration of how this affects the initial hypothesis and of the strength of additional evidence against the criteria.
- Step 3: Holistic judgement about how the totality of evidence best-fits one of the three rating descriptors.

230. The evidence that is considered in each step of assessment is displayed diagrammatically in Figure 7. Provider assessments are based primarily on provider-level evidence and subject assessments are based primarily on subject-level evidence. We will test the ways in which some subject-level information informs provider-level assessments, and vice versa, as follows:

- a. At subject level, initially considering data only about the subject to inform the initial hypothesis, and then introducing some information about the provider in step 2, alongside the subject submission.
- b. At provider level, including metrics split by subject in step 1; and then an opportunity for the Main Panel to consider the profile of subject ratings when finalising the provider ratings.

Table 10: Summary of provider- and subject-level information used in each step

Step	Subject-level assessment	Provider-level assessment
<p>Pre-assessment: Consider the context</p>	<p>Subject-level contextual data:</p> <ul style="list-style-type: none"> • profile of students studying the subject • details of courses covered. <p>Maps of subject-level geographic context.</p>	<p>Provider-level contextual data:</p> <ul style="list-style-type: none"> • profile of all students at the provider • number of students taught in each subject. <p>Maps of provider-level geographic context.</p>
<p>Step 1: Form the metrics-based initial hypothesis</p>	<p>Subject-level metrics:</p> <ul style="list-style-type: none"> • core metrics • metrics split by year and student characteristics • metrics detail sheets. 	<p>Provider-level metrics:</p> <ul style="list-style-type: none"> • core metrics • metrics split by year and student characteristics • metrics split by subject • metrics detail sheets. <p>Where applicable:</p> <ul style="list-style-type: none"> • supplementary degree attainment data and grade inflation data • additional data about part-time provision.
<p>Step 2: Consider other evidence against the criteria</p>	<p>Provider summary statement</p> <p>Subject-level submission:</p> <ul style="list-style-type: none"> • contextual information about the subject • explanation of subject metrics • additional evidence against the criteria. <p>Summary of provider-level data:</p> <ul style="list-style-type: none"> • contextual data • core metrics and splits (by year, student characteristics and subject). 	<p>Provider-level submission:</p> <ul style="list-style-type: none"> • contextual information about the provider • explanation of provider metrics • additional evidence against the criteria.
<p>Step 3: Best-fit holistic judgement</p>	<p>Totality of all the above evidence.</p>	<p>Totality of all the above evidence.</p> <p>The rating will be tested against the profile of final subject ratings.</p>

Consideration of contextual information

231. For each case, panel members should initially consider the contextual information to aid their understanding of the provider or subject being assessed, its operating context, the range of provision involved, the make-up of the student body, and the geographic context (especially in relation to employment outcomes).

232. Contextual information should not, in itself, be considered evidence of teaching excellence or student outcomes. Contextual information is intended to help the panel understand the context of the students and the provision being assessed, and aid interpretation of the evidence available in the metrics and the submission.
233. During subsequent steps of the assessment panel members may return to the contextual information to support their interpretation of metrics and consideration of the evidence presented in the submission. In their submissions, providers may wish to draw attention to particular contextual information and explain how this may affect their metrics or other evidence presented.
234. The contextual information is tailored for subject- and provider-level assessment, as set out in Figure 5.

Step 1: Review the metrics to form an initial hypothesis

235. Panel members will review the metrics data initially by applying a formula which is based only on core metric flags in the majority mode of provision (**step 1a**). This is the starting point for determining the initial hypothesis, and is based on a crucial but small part of the metrics evidence. They will then review all the metrics data, taking account of a number of other factors and will make a judgement about – rather than applying a formula to – the initial hypothesis (**step 1b**). In making this judgement they should also consider how weakly or strongly held the initial hypothesis is, and which specific elements in the data they wish to test against evidence in the submission.

Step 1a: Starting point for the initial hypothesis

236. Panel members and assessors will be presented with the starting point for the initial hypothesis, which has been calculated by the OfS by applying a formula to the core metrics flags, in the delivery mode in which the provider teaches the most students (i.e. full or part time). Where there is a similar number of students in both delivery modes, a starting point for the initial hypothesis has been calculated for both (see paragraph 270).
237. For this pilot, the formula and the associated weightings of the core metrics have been revised to take into account of the expanded basket of nine core metrics, and to allow greater differentiation in the initial stages of assessment.
238. For the purpose of calculating the starting point, the weightings of core metrics are as follows:
- a. The five core metrics based on the NSS each have a weight of 0.5.
 - b. The continuation metric has a weight of 2.0.
 - c. The three employment metrics (based on the DLHE survey and LEO dataset) each have a weight of 1.0.
239. In its 'lessons learned' exercise after TEF Year Two, the DfE determined that NSS metrics should account for one-third of the overall weighting of metrics when calculating the initial hypothesis. The weightings for this pilot have been set to maintain this overall weighting of

NSS metrics. The weighting for continuation has been increased to recognise the importance of continuation as an indicator of student success, which also underpins excellence in other areas.

240. These weightings are used to convert each core metric flag to a 'value' as set out in Table 11.

Table 11: Flag values by metric type

	+ or ++	- or --	No flag or not reportable
NSS metrics	0.5	-0.5	0
Continuation metric	2.0	-2.0	0
Employment metrics	1.0	-1.0	0

241. Table 12 sets out the formula for calculating the starting point of the initial hypothesis that we will test in the pilot. The formula is based on broadly the same proportions of positive or negative flag values associated with Gold and Bronze starting points as those used in previous TEF exercises.

242. The starting point in the pilot includes borderline ratings of Gold/Silver and Silver/Bronze, so that the assessment begins with a wider range of possible outcomes, and is then refined as all the available evidence is considered into a final rating (in step 3) of Gold, Silver or Bronze.

Table 12: Formula for calculating the step 1a initial hypothesis

Starting point	Value of core metric flags
Gold	The value of positive flags is at least 3.5, and there are no negative flags.
Gold/Silver	The value of flags is within 1.0 of a Gold starting point.
Silver	All other combinations of flag values.
Silver/Bronze	The value of flags is within 1.0 of a Bronze starting point; or the value of negative flags is at least 2.5, but the value of positive flags is greater than the value of negative flags.
Bronze	The value of negative flags is at least 2.5, and the value of positive flags is less than the value of negative flags.

243. Figure 9 sets out the starting point for each possible combination of flag values.

Figure 9: Flag combinations and their starting points

		Total value of positively flagged metrics															
		0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5
Total value of negatively flagged metrics	0	S	S	S	S	S	G/S	G/S	G	G	G	G	G	G	G	G	G
	0.5	S	S	S	S	S	S	G/S	G/S	G/S	G/S	G/S	G/S	G/S	G/S	G/S	
	1	S	S	S	S	S	S	S	G/S	G/S	G/S	G/S	G/S	G/S	G/S		
	1.5	S/B	S/B	S/B	S/B	S/B	S	S	S	S	S	S	S	S			
	2	S/B	S/B	S/B	S/B	S/B	S/B	S	S	S	S	S	S				
	2.5	B	B	B	B	B	S/B	S/B	S/B	S/B	S/B	S/B					
	3	B	B	B	B	B	B	S/B	S/B	S/B	S/B						
	3.5	B	B	B	B	B	B	B	S/B	S/B							
	4	B	B	B	B	B	B	B	B								
	4.5	B	B	B	B	B	B	B									
	5	B	B	B	B	B	B										
	5.5	B	B	B	B	B											
	6	B	B	B	B												
	6.5	B	B	B													
	7	B	B														
	7.5	B															

244. The introduction of borderline ratings of Gold/Silver and Silver/Bronze in step 1a does **not** imply that subsequent steps can only result in an adjacent rating (either side of the borderline). There is **no limit** to how far the rating in subsequent steps may move away from the step 1a initial hypothesis, regardless of whether the step 1a rating is borderline or not.

245. In all cases, the starting point for the initial hypothesis should **not be interpreted as indicative** of the final rating. It is simply a starting point for the assessment process. The core metrics provide proxies for performance across all three aspects of quality, but do not directly or fully address all 11 criteria. The final rating will be determined following consideration of all the available evidence in both the metrics and the submission, against the full set of criteria.

246. Where there are a similar number of students in both delivery modes, an initial hypothesis (step 1a) should be calculated for both modes (see paragraph 270).

Step 1b: Metrics-based initial hypothesis

247. After calculating the starting point, panel members should review all the metrics and make a judgement about the metrics-based initial hypothesis. They should consider, in particular:

- very high or low absolute values
- split metrics
- the minority mode of delivery
- other factors.

248. At provider level only, they should also consider (where applicable):

- a. Supplementary degree attainment data relating to differential attainment and to grade inflation. (This applies to classified degrees at providers with degree awarding powers.)
 - b. Additional data about part-time provision. (This applies where the provider has at least 35 per cent of its students studying part-time, and has opted to submit this additional data.)
249. The initial hypothesis, at the end of step 1b, will be arrived at as a matter of expert judgement, and is not formulaic. As in previous TEF exercises it will be on the scale of Gold, Gold/Silver, Silver, Silver/Bronze, or Bronze. In addition to determining the initial hypothesis rating on this scale, panel members should consider:
- how firmly or weakly they hold the initial hypothesis, depending on the strength of the metrics evidence
 - whether there are any specific elements in the data they wish to test against evidence in the submission.

Very high or low absolute values

250. As well as considering how a provider or subject is performing against its benchmarks, panel members should consider very high or low absolute values. The 'absolute value' is the actual value of the indicator for a given metric: for example, the actual proportion of students who are satisfied with 'Teaching in my course'. Where the absolute value is in the top or bottom 10 per cent of the sector, it will be identified as a very high or very low absolute value. It will be marked with either an asterisk (*) for very high, or an exclamation mark (!) for very low. Where a very high or low value is contradicted by a split flag, or is not statistically reliable, this will be indicated. See the 'Guide to subject-level pilot data' paragraphs 86-96 for a detailed description of how these markers are calculated and presented.²⁸
251. A **very high absolute value** suggests that a provider's performance in that metric is so high that, in absolute terms, the experience or outcome for students regarding that metric is outstanding. A very high absolute value that is statistically reliable should be interpreted as follows:
- a. If the metric is **positively flagged**, an asterisk reinforces the judgement that the provider is performing exceptionally well in this area.
 - b. If the metric is **unflagged**, and does not have negative flags in any of its split metrics, an asterisk should be considered in a similar way to a positive flag in determining the final position of the initial hypothesis.
 - c. If the metric has a **negative flag**, or has negative flags for any of its split metrics, an asterisk should not be taken into account when forming the initial hypothesis.
252. A **very low absolute value** suggests that a provider's performance in that metric is so low that, in absolute terms, the experience or outcome for many students regarding that metric

²⁸ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

is not good. A very low absolute value that is statistically reliable should be interpreted as follows:

- a. If the metric has a **negative flag**, an exclamation mark should reinforce the judgement that the provider is performing poorly in this area.
- b. If the metric is **unflagged**, and does not have positive flags in any of its split metrics, an exclamation mark should be considered in a similar way to a negative flag in determining the final position of the initial hypothesis.
- c. If the metric is **positively flagged**, or has positive flags for any of its split metrics, an exclamation mark should not be taken into account when forming the initial hypothesis.

253. Panel members will continue to be able to consider all evidence, including the presence or absence of very high and low absolute values, holistically, and to exercise their academic judgement accordingly in the subsequent steps of assessment.

Splits by student characteristics

254. Panel members should consider how the provider or subject performs with respect to different student groups. Performance with respect to certain student groups, particularly those who may be disadvantaged, must be taken into account in the initial hypothesis, and in determining the final rating. This is particularly relevant to the 'Positive outcomes for all' criterion.

255. In considering performance across student groups as part of the initial hypothesis:

- a. At subject level, panel members will consider flags in the splits by student characteristics.
- b. At provider level, panel members will consider flags in the splits by student characteristics, and any significant differences in the absolute values for different student groups.

256. When considering flags in the splits by student characteristics, no weight should be assigned to a split that does not display a flag. This is because split metrics are less likely to result in a flag than core metrics, because of their smaller cohort sizes. Panel members should focus on those split metrics that do display flags, and in particular should:

- a. Look horizontally across each metric in the workbook, and take account of split flags that differ from the core metric.
- b. Look vertically at the split flags for each student group, and take account of multiple flags for students from particular groups, which may reflect on performance in relation to criterion SO3 (Positive outcomes for all).

257. Panel members should consider whether and how far these split flags influence the initial hypothesis, either to a higher or lower position, or to hold their initial hypothesis more strongly or weakly, and whether there are particular student groups about whom they would wish to see further evidence in the submission.

258. At provider level only, panel members should also consider significant differences in the absolute values between student groups. This information is not benchmarked and should only tentatively influence the initial hypothesis, pending consideration of any explanation provided in the submission about differential outcomes. We will explore with the panels how far this data is useful in considering differential outcomes across student groups, and how the data could be developed further beyond the pilot.

Splits by subject (at provider level)

259. When forming the initial hypothesis about a provider, panel members should consider the profile of flags across the subjects taught by the provider. This may provide an indication of how far metrics-based performance is consistent across the provider, or may identify particular subjects that contribute in particular to core flags or splits by student characteristics.

260. However, subject splits typically involve smaller cohort sizes than a provider's core metrics, and panel members should not draw conclusions from unflagged subject splits. For example, if a core metric is positively flagged and the subject splits are a mix of positive and unflagged, this does not indicate a lack of consistently strong performance across those subjects.

261. We will test in the pilot how far the subject splits are useful in informing provider-level assessments, and if appropriate develop further guidelines on how they can be interpreted and used.

Splits by year

262. The core metrics aggregate data over one, two or three years depending on how much data is available. The metrics should be considered as a 'snapshot' of performance over this period. The metrics are additionally presented as split by year; however, these year splits should not be considered as a trajectory, and the presence or absence of a trend should not affect the initial hypothesis. The year splits are provided to aid understanding of how many years of reportable data are captured in each core metric. Flags in the year splits should inform the initial hypothesis **only where one of the following applies**:

- a. The core metric is not reportable when aggregating all years of available data. In these rare cases, any flags in the year splits should be regarded as a substitute for the core metric flags when calculating the starting point (step 1a).
- b. The core metric is neutral and there are one or more flags in the year splits. In these cases the flags should be considered as an indication of positive or negative performance for part of the period under consideration, but not (in the absence of a core flag) for the whole period.

263. In step 2 when considering the submission, panel members should seek evidence in particular about the **impact** of policies, practices or initiatives. Where the submission claims that initiatives have had a recent impact, panel members may refer to the year splits in the relevant metrics as potential corroboration for these claims. Providers may wish to draw attention in their submissions to any relevant year splits for this purpose.

Additional data about part-time provision

264. Where the provider has at least 35 per cent of its students studying part-time and has included the additional page of quantitative information, this should be considered in step 1b alongside the part-time metrics. Any data that does not relate specifically to the following should be disregarded:
 - a. Student satisfaction, for part-time students who are not in scope of the NSS (that is, students on courses of 1 full-time equivalent or less, such as HNCs).
 - b. Continuation, for part-time students who are studying flexibly.
 - c. Employment, for part-time students already in work or not seeking work after their studies.
265. The greater the robustness of the additional data provided, and the more the information is put in the context of national comparators, the more weight may be placed on it.

Additional factors in reviewing the metrics

266. Panel members should consider a number of additional factors related to the interpretation of the core and split metrics in order to refine the initial hypothesis. These are:
 - a. **Distribution of flags:** In addition to the number of flags and very high or low absolute values, panel members should consider how they are distributed across the three aspects of quality. If they are concentrated in – or absent from – one or more aspects, that may influence the judgement.
 - b. **Differences from benchmarks:** In addition to considering whether a metric is flagged, panel members should consider how far the metrics differ from their benchmarks, and the statistical significance of those differences. Table 13 sets out how these should be interpreted. For the purpose of the pilot this includes consideration of metrics with Z-scores between 1.65 and 1.96 as **potential** indicators of performance, and we will explore whether this could help maximise the use of available data at subject level. We will seek specific feedback from pilot participants and panels on the robustness and utility of this approach.

Table 13: Interpretation of Z-scores and differences from benchmark

Z-score	Likelihood of the difference from benchmark being due to chance	How this the difference from benchmark should be interpreted
Below 1.65	More than 10 per cent	The difference from benchmark should be disregarded as it lacks statistical significance.
1.65 to 1.96	5 to 10 per cent	A difference of at least two percentage points should tentatively be regarded as a potential indicator of performance, to be tested against evidence in the submission.
1.96 to 3	5 per cent or less	A difference of at least two percentage points will be flagged and regarded as an indicator of increasingly strong or weak performance, the greater the difference is.
3 or above	Negligible (Note that Z-scores of any value of 3 or above will for practical purposes be considered as equivalent)	A difference of between two and three percentage points will be flagged and regarded as an indicator of performance. A difference of at least three percentage points will be double flagged and regarded as an indicator of increasingly strong or weak performance, the greater the difference is.

- c. **Limitations in the data:** Panel members should consider limitations or other features of the data that would lead the initial hypothesis to be more weakly held, and where further information in the submission would be needed to inform a judgement. This is likely where:
- i. There are no or very few flags, because of very small cohorts.
 - ii. There is a core metric with a contrary flag in the split metric.
 - iii. There are a mixture of positive and negative flags.
 - iv. There is a concentration of positive or negative flags in one or more aspects that is not replicated in other aspects.
 - v. There are unreportable metrics (either for the majority mode, or for the minority which represents a significant proportion of the provider's students).
 - vi. There are fewer than three years of core metrics.
 - vii. The provider or subject's own data dominates its benchmarks.

viii. There is a high proportion of local students.

Minority mode of delivery

267. The performance of the provider or subject in each delivery mode (i.e. full time or part time) should be considered in proportion to the number of students in each mode.
268. All the metrics in the minority mode – including the core metric flags, splits, very high or very low absolute values and other factors – should be taken into account in a similar way as for the majority mode, and should affect the initial hypothesis in proportion to the number of students (headcount) in the minority mode.
269. Where this proportion is fewer than 35 per cent of students, all the minority mode metrics should be considered within step 1b.

Similar numbers of students in each mode

270. Where the proportions of full-time and part-time students are both greater than 35 per cent²⁹ – that is, there are a similar number of students in each delivery mode – an initial hypothesis should be considered separately for each mode, and then combined at the end of step 1b, as follows:
 - a. Panel members should calculate a starting point separately for each mode.
 - b. Panel members should separately consider the remaining metrics and arrive at a step 1b initial hypothesis for each mode.
 - c. Panel members should combine these into a single, final, initial hypothesis at the end of step 1b, taking into account the relative numbers of students in each mode.

Supplementary degree attainment data (at provider level only)

271. Where the supplementary degree attainment data is available at provider level, panel members should initially consider, in step 1b:
 - a. Whether there are any significant gaps in degree attainment between students from disadvantaged and other backgrounds; and if so whether they have narrowed over the six-year period covered by the data, and how they relate to the core metric splits by student characteristics.
 - b. Whether there has been potential inflation in the proportion of 1sts and 2:1s over the 10-year period covered by the data; and if so how this might relate to changes in students' entry qualifications for the most recent six-year period.
272. In the pilot the supplementary attainment data is not benchmarked and will not be flagged. While they should for the purposes of the pilot be considered initially in step 1b, they should only tentatively influence the initial hypothesis pending consideration of relevant information in the submission. In particular:

²⁹ The metrics workbooks will clearly indicate whether this applies.

- a. If the degree attainment data provides a potential indication of performance, this should be held tentatively in step 1b and reconsidered in step 2 alongside any information relating to 'Positive outcomes for all', including any explanation of how attainment gaps are identified and addressed.
- b. If the grade inflation data provides a potential indication of performance, this should be held tentatively in step 1b and reconsidered in step 2 alongside any information relating to 'Rigour and stretch', including any explanation of how standards are maintained and what factors have contributed to any change in the profile of degree classifications.

273. Through the pilot we will explore this data further and how it might be developed as potential indicators of excellence. We will also explore further how best to make use of 'supplementary data' alongside metrics and submissions.

Strength of the initial hypothesis

274. Step 1b must result in an initial hypothesis of Gold, Silver or Bronze, or a borderline rating between these. Panel members should also explicitly consider:

- a. How firmly or weakly they hold the initial hypothesis, depending on the overall strength of the metrics evidence.
- b. Whether any specific aspects of the data are considered as no more than tentative indications of performance, until tested against evidence in the submission. This applies to, but is not limited to:
 - i. The supplementary degree attainment data, in relation to both differential attainment and grade inflation (at provider level only).
 - ii. Unflagged metrics where the difference from benchmark is at least two percentage points and the z-score is between 1.65 and 1.96.

275. A Gold (or Bronze) initial hypothesis should be more firmly held where the metrics data is more consistently positive (or negative) across the core flags and other considerations described above. A Silver starting point should be more firmly held where the metrics indicate with a degree of confidence that performance is consistently close to benchmark.

276. Where an initial hypothesis is firmly held, panel members must still fully consider the totality of evidence across the metrics and submission, and make a best-fit judgement against the rating descriptors.

277. The more weakly the initial hypothesis is held, the more heavily panel members should rely on evidence in the submission to determine a best-fit rating. This includes cases where the initial hypothesis 'defaults' to Silver based on an absence of flags for metrics that are unreportable or cover very small cohorts. During the pilot we will work with providers on how submissions might compensate for such limitations in the data.

278. After reviewing the metrics and forming an initial hypothesis, panel members will review the submission (step 2) before making a holistic best-fit judgement (step 3).

Step 2: Review of the submission

279. In reviewing the provider-level or subject-level submission, panel members should consider:
- a. Any further contextual information it provides.
 - b. The strength of evidence of excellence that it provides against all three aspects of quality, and in particular how far the submission demonstrates the impact or effectiveness of policies, practices or initiatives.
 - c. How the submission engages with and explains metrics-based performance, including potential mitigation of negative flags, and any alternative evidence provided to addresses any unreportable metrics.
 - d. How far student views have been incorporated into the submission, and provide panel members with confidence that the submission reflects student views at the provider. Where students have explicitly chosen not to engage with the submission, however, providers should not be disadvantaged (please see paragraphs 176-185).
280. All providers taking part in the TEF will already have met the high provider-level baseline quality thresholds for the UK sector. Where **neither the metrics nor the submission** provide evidence about a particular criterion, panel members should assume that performance in relation to that criterion is at baseline level. This may include areas where the metrics are unreportable or lack statistical significance (typically due to very small cohorts), and the submission has not provided evidence. The onus is on the provider to demonstrate that performance exceeds the baseline or is outstanding.
281. For additional evidence to improve the initial hypothesis, panel members should expect to see clear, well supported evidence that aligns with the relevant rating descriptor across the three aspects of quality. In particular, for providers to achieve a Silver or Gold rating, panel members should look to see clear evidence, from the metrics in combination with the additional evidence, of excellent or outstanding performance against all three aspects of teaching excellence.
282. As in step 1, weight commensurate with the proportion of students in each mode should be placed on evidence relating to that mode.
283. Panel members should give no weight to evidence that is not relevant to the criteria.
284. Where a submission suggests that any elements of the metrics or the provider's performance can be explained by the fact that a provider has high absolute values, panel members should only consider such arguments to be valid if the relevant metric has an asterisk. If the relevant metric does not have an asterisk, panel members should not consider as valid assertions in submissions that make reference to high absolute values.

Main Panel's consideration of differential outcomes and grade inflation

285. At provider level, Main Panel members should specifically consider evidence in the submission that relates to the supplementary degree attainment data (where applicable):
- a. How the submission explains the data relating to differential attainment, and any associated evidence about how the provider identifies and addresses potential differential outcome and supports positive outcomes for all groups of students.
 - b. How the submission explains the data relating to grade inflation, including evidence that any substantial increases in the proportions of 1sts and 2:1s awarded have occurred for legitimate reasons, and broader evidence about rigour and stretch at that provider.

Subject panels' consideration of provider-level information

286. We will test the following approach to taking account of provider-level information as part of subject-level assessments:
- a. The initial hypothesis will be based on metrics data purely at subject level.
 - b. In step 2, subject panel members should consider some provider-level information as well as the subject-level submission. This is intended to enable subject-level assessments to take account of the institutional context, policies and facilities that impact on provision at subject level. We are also testing whether it might help mitigate certain limitations in subject-level data.
287. In step 2, subject panels members should:
- a. Consider the provider summary statement, to form an understanding of the institutional context, policies and facilities that may impact on provision at subject level.
 - b. Review the subject submission and the subject-specific evidence it contains, according to the guidance in 141-185.
 - c. Refer to the summary of provider-level data. This data comprises the provider-level contextual data and the provider-level core and split metrics (including subject splits). We are testing whether this data is useful as evidence to inform subject-level assessments, for example:
 - i. To accompany the provider summary statement which may refer to provider-level metrics
 - ii. As a potential means of mitigating limitations in subject-level data. Where there are non-reportable metrics or unflagged metrics with very small cohorts, subject panel members should consider how far the provider-level data might provide an alternative indicator of performance, when considered alongside the submission.
288. We will seek subject panel members' feedback on the usefulness of provider-level information as part of subject-level assessments.

Step 3: Best-fit holistic judgement

289. Having reviewed the metrics and submission, panel members should consider the totality of evidence available, to make a 'best-fit' judgement against one of the three rating descriptors (as shown in Table 4).
290. The rating descriptors set out typical characteristics at each level of excellence, related to the criteria. The rating descriptors are tailored at provider and subject levels. Panel members should use these descriptors to make a best-fit judgement.
291. The evidence about a provider or subject need not align with all components of a descriptor to receive that rating, and may well align with some components of a different descriptor. Panel members should make an overall best-fit judgement.
292. In doing so, there is no set weighting between the metrics and the submission. The totality of evidence should be considered, and the relative strength of evidence available in the metrics and submission may vary. If the metrics do not provide clear-cut or consistent evidence, the submission will be especially important. Conversely, if the submission includes limited evidence, proportionately more weight should be placed on the metrics – in this case the available evidence (that is, primarily the metrics) should be considered against the rating descriptors, and the final rating may differ from the initial hypothesis. In particular:
- a. Where the initial hypothesis is Gold (or Silver), the panel should expect to see additional evidence in the submission that supports that hypothesis, to merit a best-fit judgement of Gold (or Silver) based on the totality of evidence.
 - b. Where the initial hypothesis 'defaults' to Silver largely due to limitations in the data (such as non-reportable metrics or a lack of flags due to small cohorts) the panel should expect substantive additional evidence in the submission that aligns with the Silver (or Gold) rating descriptor, in order to merit a best-fit judgement of Silver (or Gold).
293. At the final stage of the assessment, after all of the subject ratings have been determined, the Main Panel will have the opportunity to consider the profile of subject ratings when finalising the provider ratings. We anticipate that this may affect the holistic judgement either where the rating based on provider-level information is very close to a borderline, or where the profile of subject ratings appears to contradict the provider rating. In the latter case, the Main Panel may wish to reconsider the provider rating, taking into account all of the provider-level evidence and the profile of subject ratings. We will test in the pilot whether this is an appropriate way to achieve coherence between the subject-level and provider-level ratings.

Outcomes

Ratings and statements of findings

294. Each provider that participates in the pilot will receive indicative ratings and statements of findings at provider level, and for all of their assessed subjects. These will not be published.

295. Statements of findings will set out at a high level the rationale for the rating. We will explore how the statements could be made most useful to students and providers, and support enhancement, in a way that is scalable for full subject-level TEF.
296. In addition to providing statements of findings for individual ratings we intend to further explore and test ways that TEF can support sector-wide enhancement. For a sample of their subjects, we will ask panel members to produce reports identifying good practice and areas for development, based on a synthesis of the evidence submitted in the pilot. These reports may be published, as well as the panel's reports that provide feedback on the assessment process.

Communication and use of outcomes

297. In full subject-level TEF we envisage publishing each provider's and subject's metrics, submissions, ratings and statements of findings as official sources of information for students. As the pilot is developmental, we will not publish this information from the pilot.
298. In the interests of transparency, and to enable preparation for full subject-level TEF, we will publish summary sector-level data from the subject-level TEF pilot. This will not be attributable to any provider or individual subject. We also expect participating providers to share outcomes only within their institution, including with their student representative bodies.
299. Alongside the pilot, we will commission user testing of subject-level TEF outcomes with applicants. This activity will test applicants' reactions to subject- and provider-level TEF outcomes and inform us how they can most usefully be presented in future, alongside other information sources.

Evaluation of the pilots and next steps

300. The pilot will be evaluated against the themes set out in paragraphs 25. Pilot participants, TEF student representatives and panel members will be expected to fully contribute to evaluation activities, including evaluation activities undertaken by consultants commissioned by the OfS.
301. Full evaluation findings and aggregate data will be published at the conclusion of the pilot. The findings will be considered alongside the outcomes from the government's independent review of the TEF to inform the final design and implementation of full subject-level TEF.

List of abbreviations

CAH	Common Aggregation Hierarchy
DfE	Department for Education
DLHE	Destination for Leavers of Higher Education survey
HERA	Higher Education and Research Act
HESA	Higher Education Statistics Agency
JACS	Joint Academic Coding System
LE	Learning environment
LEO	Longitudinal Education Outcomes
NSS	National Student Survey
OfS	Office for Students
POLAR	Participation of local areas
PSRB	Professional, statutory and regulatory bodies
SO	Student outcomes and learning gain
TEF	Teaching Excellence and Student Outcomes Framework
TQ	Teaching quality
WP	Widening participation

Annex A: TEF Subjects list – mapping

Pilot Subject Panel	CAH2 subjects to be assessed (adapted for the TEF pilot)	Sub-disciplines which comprise the subject (identified by CAH3 subject classifications)
Medical sciences	Medicine and dentistry	medical sciences (non-specific)
		medicine (non-specific)
		medicine by specialism
		dentistry
	Pharmacology, toxicology and pharmacy	pharmacology
		toxicology
		pharmacy
	Medical sciences	medical technology
		healthcare science (non-specific)
		biomedical sciences (non-specific)
		anatomy, physiology and pathology
	Psychology	others in subjects allied to medicine (medical sciences)*
		psychology (non-specific)
		applied psychology
		developmental psychology
psychology and health		
Veterinary sciences	others in psychology	
	veterinary medicine and dentistry	
Nursing and allied health subjects	Veterinary sciences	others in veterinary sciences
		nursing (non-specific)
	Nursing and midwifery	adult nursing
		community nursing
		midwifery
		children's nursing
		dental nursing
		mental health nursing
		learning disabilities nursing
		others in nursing
	Allied health	health sciences (non-specific)
		nutrition and dietetics
		ophthalmics
		environmental and public health
		physiotherapy
complementary and alternative medicine		
counselling, psychotherapy and occupational therapy		
others in subjects allied to medicine (allied health)*		
Sport and exercise sciences	sport and exercise sciences	

* The CAH3 'Others in subjects allied to medicine' has been disaggregated at the level of JACS code, broadly in line with the occupational aspects of the subject. JACS codes relating to biomedical sciences have been mapped to the 'Medical Sciences' CAH2, while those relating more closely to allied health professions have been mapped to the 'Allied health' CAH2. Full details can be found in the TEF metrics technical document.

Pilot Subject Panel	CAH2 subjects to be assessed (adapted for the TEF pilot)	Sub-disciplines which comprise the subject (identified by CAH3 subject classifications)
Natural sciences	Biosciences	biosciences (non-specific)
		biology (non-specific)
		ecology and environmental biology
		microbiology and cell science
		plant sciences
		zoology
		genetics
		molecular biology, biophysics and biochemistry
		hair and beauty sciences
		others in biosciences
	Physics and astronomy	physics
		astronomy
	Chemistry	chemistry
	General, applied and forensic sciences	physical sciences (non-specific)
		forensic and archaeological sciences
		sciences (non-specific)
		natural sciences (non-specific)
	Mathematical sciences	mathematics
		operational research
statistics		
Engineering and technology	Engineering	engineering (non-specific)
		mechanical engineering
		production and manufacturing engineering
		aeronautical and aerospace engineering
		naval architecture
		bioengineering, medical and biomedical engineering
		civil engineering
		electrical and electronic engineering
		chemical, process and energy engineering
		others in engineering
	Materials and technology	materials science
		minerals technology
		materials technology
		polymers and textiles
		maritime technology
		biotechnology
	others in technology	
	Computing	computer science
		information technology
information systems		
software engineering		
artificial intelligence		
computer games and animation		
business computing		
others in computing		

Pilot Subject Panel	CAH2 subjects to be assessed (adapted for the TEF pilot)	Sub-disciplines which comprise the subject (identified by CAH3 subject classifications)
Natural and built environment	Agriculture, food and related studies	animal science
		agricultural sciences
		agriculture
		rural estate management
		forestry and arboriculture
		food and beverage studies (non-specific)
		food sciences
	food and beverage production	
	Geography, earth and environmental studies	earth sciences
		geography (non-specific)
physical geographical sciences		
human geography		
environmental sciences		
Architecture, building and planning	others in geographical studies	
	architecture	
	building	
	landscape design	
	planning (urban, rural and regional)	
Social sciences	Sociology, social policy and anthropology	social sciences (non-specific)
		sociology
		social policy
		anthropology
		development studies
		cultural studies
	Economics	economics
Politics	politics	

Pilot Subject Panel	CAH2 subjects to be assessed (adapted for the TEF pilot)	Sub-disciplines which comprise the subject (identified by CAH3 subject classifications)
Business and law	Law	law
	Business and management	business and management (non-specific)
		business studies
		marketing
		management studies
		human resource management
		tourism, transport and travel
		finance
		accounting
		others in business and management
Education and social care	Education and teaching	education
		teacher training
	Health and social care	social work
		childhood and youth studies
		health studies

Pilot Subject Panel	CAH2 subjects to be assessed (adapted for the TEF pilot)	Sub-disciplines which comprise the subject (identified by CAH3 subject classifications)
Arts	Creative arts and design	creative arts and design (non-specific)
		art
		design studies
		cinematics and photography
	Performing arts	others in creative arts and design
		performing arts (non-specific)
		music
	Media, journalism and communications	drama
		dance
information services		
publicity studies		
Humanities	English studies	publishing
		journalism
		media studies
		English studies (non-specific)
		English language
		literature in English
	Languages and area studies	studies of specific authors
		creative writing
		others in English studies
		linguistics
Celtic studies (non-specific)		
Gaelic studies		
Welsh studies		
others in Celtic studies		
French studies		
German and Scandinavian studies		
History and archaeology	Italian studies	
	Iberian studies	
	Slavic studies	
	Asian studies	
	African and modern Middle Eastern studies	
Philosophy and religious studies	American and Australasian studies	
	others in language and area studies	
Combined and general studies	history	
	history of art, architecture and design	
	archaeology	
	heritage studies	
Combined and general studies	classics	
	philosophy	
	theology and religious studies	
	humanities (non-specific)	
Combined and general studies	liberal arts (non-specific)	
	combined, general or negotiated studies	
	personal development	

Annex B: TEF subject-level pilot – full metrics descriptions

Metrics based on the National Student Survey

1. Paragraphs 2 to 24 of this annex provide a full description of the metrics to be used in the 2018-19 subject-level TEF pilot which are based on the National Student Survey (NSS). Unless otherwise stated, the details of the metrics' coverage, exclusions and benchmarking factors (described in paragraphs 4 to 10) are applicable to all five of the metrics derived using NSS responses.
2. The NSS asks a range of individual questions which are organised into different sets. These are known as NSS question scales, with each representing a different theme. In responding to an individual NSS question, students indicate their agreement with each statement on a five-point scale. Across the questions that make up a given scale, total agreement by each student is calculated as the percentage of responses that are 'Agree' or 'Strongly agree'. Questions marked with 'N/A' or not answered are ignored. An example, based on the 'Teaching on my course' scale (which comprises Questions 1 to 4), is given in Table B1.

Table B1: Example data

Student	Q1	Q2	Q3	Q4	Percentage agree
A	Strongly agree	Strongly agree	Agree	Neither	75
B	Strongly agree	Strongly agree	Agree	N/A	100
C	Strongly agree	Agree	Agree	Disagree	75
D	Strongly agree	Agree	Disagree	Disagree	50
E	Agree	Disagree	N/A	N/A	50

3. In this example, the total percentage agreement for the provider would be 70 per cent (the sum of percentages divided by the number of students: $(75+100+75+50+50= 350) \div 5 = 70$).

Coverage (applicable to all NSS-based metrics)

4. The NSS is targeted at all final year undergraduates. Non-final year students on flexible provision or who change their study plans may also be included by participating providers.
5. The NSS covers UK, other EU and non-EU students.
6. The NSS-based metrics cover students surveyed during the spring of:
 - 2016 (defines the Year 1 year-split metric)
 - 2017 (defines the Year 2 year-split metric)
 - 2018 (defines the Year 3 year-split metric).
7. The NSS response rate in 2018 was 70 per cent.

Exclusions (applicable to all NSS-based metrics)

8. The following exclusions apply:

- a. Students who did not reach the final year of their course in any of the three most recent years.
- b. Students not aiming for an undergraduate level qualification, or aiming for a qualification of 1 FTE or lower.
- c. Students who were not on a full-time or part-time mode of study.

Benchmarking factors (applicable to all NSS-based metrics)

9. The factors used to benchmark the NSS-based metrics for full-time students are:

- subject of study
- age of entry
- ethnicity
- disability
- level of study
- year of cohort.

10. The factors used to benchmark the NSS-based metrics for part-time students are:

- subject of study
- age of entry
- disability
- year of cohort.

'Teaching on my course' metric

11. This metric is based on student's responses to NSS questions 1 to 4 which cover the NSS scale 'Teaching on my course'.

12. In 2016 these questions were as follows:

- Q1** – Staff are good at explaining things.
- Q2** – Staff have made the subject interesting.
- Q3** – Staff are enthusiastic about what they are teaching.
- Q4** – The course is intellectually stimulating.

13. In 2017 and 2018 these questions were as follows:

- Q1** – Staff are good at explaining things.
- Q2** – Staff have made the subject interesting.

Q3 – The course is intellectually stimulating.

Q4 – My course has challenged me to achieve my best work.

‘Assessment and feedback’ metric

14. This metric is based on student’s responses to NSS questions which cover the NSS scale ‘Assessment and feedback’.

15. In 2016 these were Questions 5 to 9, as follows:

Q5 – The criteria used in marking have been clear in advance.

Q6 – Assessment arrangements and marking have been fair.

Q7 – Feedback on my work has been prompt.

Q8 – I have received detailed comments on my work.

Q9 – Feedback on my work has helped me clarify things I did not understand.

16. In 2017 and 2018 these were Questions 8 to 11, as follows:

Q8 – The criteria used in marking have been clear in advance.

Q9 – Marking and assessment has been fair.

Q10 – Feedback on my work has been timely.

Q11 – I have received helpful comments on my work.

‘Student voice’ metric

17. This metric is based on student’s responses to NSS questions 23 to 25, which cover the NSS scale ‘Student voice’. As these questions were only present in the 2017 and 2018 NSS, there will be no ‘Student voice’ metric for ‘Year 1’, which corresponds to the 2016 NSS.

18. In 2017 and 2018 the ‘Student voice’ questions were as follows:

Q23 – I have had the right opportunities to provide feedback on my course.

Q24 – Staff value students’ views and opinions about the course.

Q25 – It is clear how students’ feedback on the course has been acted on.

19. Question 26 from this scale is not included in the calculation of the ‘Student voice’ metric.

‘Academic support’ metric

20. This metric is based on student’s responses to NSS questions which cover the NSS scale ‘Academic support’.

21. In 2016 these were Questions 10 to 12, as follows:

Q10 – I have received sufficient advice and support with my studies.

Q11 – I have been able to contact staff when I needed to.

Q12 – Good advice was available when I needed to make study choices.

22. In 2017 and 2018 these were Questions 12 to 14, as follows:

Q12 – I have been able to contact staff when I needed to.

Q13 – I have received sufficient advice and guidance in relation to my course.

Q14 – Good advice was available when I needed to make study choices

‘Learning resources’ metric

23. This metric is based on student’s responses to NSS questions which cover the NSS scale ‘Learning resources’.

24. In 2017 and 2018 the ‘Learning resources’ questions were as follows.

Q18 – The IT resources and facilities provided have supported my learning well.

Q19 – The library resources (e.g. books, online services and learning spaces) have supported my learning well.

Q20 – I have been able to access course-specific resources (e.g. equipment, facilities, software, collections) when I needed to.

Continuation metrics

25. The continuation metrics described at paragraphs 26 to 34 are based on the individualised student data captured in the Higher Education Statistics Agency (HESA) and Individualised Learner Record (ILR) student records.

26. Full-time and part-time continuation metrics are calculated by slightly different methods. The full-time metric tracks students from the date they enter a higher education provider to their activity a year later. The part-time metric tracks students from entry to their activity two years later. The full-time continuation metric is based on student activity on a census date one year and 14 days after their commencement date, and two years and 14 days for the part-time metric. Students who qualify at undergraduate or postgraduate level on or before the census date, continue at the same provider on the census date, or are studying at higher education level at another provider on the census date are deemed to have continued. All other students are deemed non-continuers.

27. To align with the census date period, an entrant year cohort is defined based on those students starting courses between the dates of 18 July and the following 17 July. This allows the activity of all students in this cohort on their census date to be determined in the following data reporting period for the full-time metric, or the data reporting period two years afterwards for the part-time metric.

28. To be counted as continuing, the student must either have qualified or be recorded as actively studying on a higher education course in the relevant HESA or ILR dataset. Students who transfer to a provider that does not submit data to HESA or ILR will be counted as non-continuers.

Coverage of the continuation metric

29. This metric includes all UK-domiciled students who are included in one of the relevant HESA or ILR datasets and registered as entrants on higher education Level 4, 5 and 6 programmes³⁰.

30. The full-time continuation metric covers students entering higher education:

- between 18 July 2013 and 17 July 2014 (defines the Year 1 year-split metric)
- between 18 July 2014 and 17 July 2015 (defines the Year 2 year-split metric)
- between 18 July 2015 and 17 July 2016 (defines the Year 3 year-split metric).

31. The part-time continuation metric covers students entering higher education:

- between 18 July 2012 and 17 July 2013 (defines the Year 1 year-split metric)
- between 18 July 2013 and 17 July 2014 (defines the Year 2 year-split metric)
- between 18 July 2014 and 17 July 2015 (defines the Year 3 year-split metric).

Exclusions of the continuation metric

32. The following exclusions apply:

- a. EU and non-EU international students.
- b. Students not registered on a first degree or other undergraduate course.
- c. Students registered at the same provider studying at the same level in the year prior to entry.
- d. Students recorded in another provider's HESA or ILR data for the same activity.
- e. Students with more than one record at a provider with the same mode and level of study.

Benchmarking factors for the full-time continuation metric

33. The factors used to benchmark the full-time continuation metric are:

- subject of study
- entry qualifications
- age on entry
- ethnicity
- Participation of Local Areas (POLAR4) quintile
- level of study.

³⁰ Throughout this annex references to Level 6 qualifications should be taken to include integrated masters' qualifications.

Benchmarking factors for the part-time continuation metric

34. The factors used to benchmark the part-time continuation metric are:

- subject of study
- age on entry
- ethnicity
- POLAR4 quintile
- level of study.

Employment metrics

‘Highly skilled employment or higher study’ metric

35. Paragraphs 36 to 45 provide a full description of this metric, which is based on the Destinations of Leavers in Higher Education (DLHE) survey.

36. This metric expresses the number of UK-domiciled leavers who say they are in highly skilled employment or studying at a higher level (or both), as a percentage of all those who are working or studying or seeking work at approximately six months after leaving. All other categories are excluded from the denominator for this indicator.

37. Leavers are asked to indicate their current activity, selecting from eight categories. They are then asked to indicate the most important activity.

38. Those who indicate they are in employment are asked to provide further detail about that employment including a job title. That job title is mapped to the Standard Occupational Classification (SOC)³¹. For this metric, jobs that are coded in SOC major groups 1 to 3 are counted as highly skilled.

39. Those who indicate they are in further study are asked to provide further detail about the type of qualification they are aiming for. This is used to determine whether this further study was at a higher level than their original study. For students who qualified at ‘Other UG’ level, the responses that are classified at a higher level are ‘higher degree, mainly by research’, ‘higher degree, mainly by taught course’, ‘postgraduate diploma or certificate’ and ‘first degree’. For all other students, the responses that are classified at a higher level are ‘higher degree, mainly by research’, ‘higher degree, mainly by taught course’ and ‘postgraduate diploma or certificate’

40. The indicator is therefore those leavers in categories 01 to 06 (where employment is in SOC 1-3, and further study is at a higher level) divided by those leavers in categories 01 to 08.

41. The full list of responses which count positively towards this metric are provided in Annex B of the ‘Guide to subject-level pilot data’³². Further detail can be found on the HESA website³³.

³¹ See www.hesa.ac.uk/support/documentation/industrial-occupational.

³² See www.officeforstudents.org.uk/advice-and-guidance/teaching/piloting-tef-at-subject-level/.

³³ See www.hesa.ac.uk/pis/emp.

Coverage of the highly-skilled employment or higher study metric

42. This metric includes all UK-domiciled leavers who are included in the relevant HESA and ILR datasets and have been awarded full Level 4, 5 or 6 qualifications.

43. The metric covers students leaving higher education in academic years:

- 2014-15 (defines the Year 1 year-split metric)
- 2015-16 (defines the Year 2 year-split metric)
- 2016-17 (defines the Year 3 year-split metric).

Exclusions of the highly-skilled employment or higher study metric

44. The following exclusions apply:

- a. EU and non-EU international students.
- b. Students who are not counted in the DLHE target population.
- c. Students who were not awarded an undergraduate Level 4, 5 or 6 qualification.
- d. Students who are recorded in another provider's ILR data for the same activity.

Benchmarking factors for the highly-skilled employment or higher study metric

45. The factors used to benchmark the highly-skilled employment or higher study metric are:

- subject of study
- entry qualifications
- age on entry
- ethnicity
- sex
- disability
- POLAR4 quintile
- level of study.

'Sustained employment or further study' metric

46. Paragraphs 47 to 64 provide a full description of the metrics to be used within the 2018-19 subject-level TEF pilot which are based on the Longitudinal Employment Outcomes (LEO) dataset, which links higher education and tax data together to chart the transition of graduates from higher education to the workplace. The LEO dataset links information about students, including their personal characteristics, their education (including schools, colleges and higher education providers attended as well as courses taken and qualifications achieved), their employment and income, and any benefits claimed.

47. This employment metric considers the proportion of qualifiers in sustained employment or further study three years after graduation, based on the LEO dataset.

48. The Department for Education's definition of sustained employment is consistent with definitions used for 16-19 accountability, and the outcome-based success measures published for adult further education. The definition of sustained employment three years after graduation looks at employment activity captured by Pay As You Earn (PAYE) employment records for at least one day a month in five out of six months in the October to March periods in 2013-14, 2014-15 and 2015-16, for graduates in 2009-10, 2010-11 and 2011-12 respectively. If they are employed in all five months from October to February, but do not have a PAYE employment record for March, then graduates must have an additional employment record in April to be considered as being in sustained employment. The definition also includes any graduates returning a self-assessment tax form in 2013-14, 2014-15 or 2015-16 stating that they have received income from self-employment during that tax year.

49. 'Sustained employment' is taken together with further study in the definition of TEF metrics: the numerator counts all graduates identified as being in sustained employment or further study. A graduate is defined as being in further study if they have a valid higher education study record at any UK publicly funded higher education institution or any English alternative provider or further education college in the HESA and ILR student records in the 2013-14, 2014-15 or 2015-16 tax year. The further study does can be a higher education qualification at any level (undergraduate or postgraduate).

50. Further detail can be found on the Gov.uk website³⁴.

Coverage of the sustained employment or further study metric

51. This metric includes all UK-domiciled leavers who are included in the relevant HESA and ILR datasets and have been awarded full Level 4, 5 or 6 qualifications.

52. The metric covers students leaving higher education in academic years:

- 2009-10 (defines the Year 1 year-split metric)
- 2010-11 (defines the Year 2 year-split metric)
- 2011-12 (defines the Year 3 year-split metric).

Exclusions of the sustained employment or further study metric

53. The following exclusions apply:

- a. EU and non-EU international students.
- b. Students who were not awarded an undergraduate Level 4, 5 or 6 qualification.
- c. Students who are recorded in another provider's ILR data for the same activity.
- d. Students at providers in Northern Ireland, or at providers that were not required to return HESA or ILR data in 2009-10, 2010-11 or 2011-12.
- e. Students without an HM Revenues and Customs (HMRC) or Department for Work and Pensions (DWP) tax or benefits record in any year, and the small minority of students

³⁴ See www.gov.uk/government/statistics/graduate-outcomes-by-degree-subject-and-university.

whose personal characteristics data was not sufficiently complete to enable linking to HMRC or DWP records.

Benchmarking factors for the sustained employment or further study metric

54. The factors used to benchmark the full-time sustained employment or further study metric are:

- subject of study
- entry qualifications
- age on entry
- ethnicity
- sex
- level of study.

55. The factors used to benchmark the part-time sustained employment or further study metric are:

- subject of study
- ethnicity
- level of study.

‘Above median earnings or in higher study’ metric

56. This employment metric is based on the LEO dataset, and considers the proportion of qualifiers in sustained employment who are earning over the median salary for 25- to 29-year-olds, or are in further study at a higher level than their original qualification.

57. Qualifiers in sustained employment or further study are defined as in the description of the ‘Sustained employment or further study’ metric. It is the population of graduates who were in sustained employment (with known salary information) or in further study who form the denominator of this metric. The numerator includes graduates within this denominator population whose PAYE or self-assessment earnings exceed the median salary for 25- to 29-year-olds in the relevant tax year three years after graduation. This is defined according to the statistics published in the annual HMRC publication series ‘Personal income statistics’³⁵. The median salaries for 25- to 29-year-olds, as published, were £20,000 in 2013-14, £21,000 in 2014-15 and £21,500 in 2015-16.

58. Earnings from PAYE are annualised by calculating the individual’s average daily wage for the days recorded in employment and multiplying this by the number of days in the tax year. Self-assessed earnings are not annualised as this calculation is not possible. This means that the metric considers raw earnings data for graduates who have only been matched to a self-assessment tax return. For a graduate with a combination of earnings from PAYE and self-assessment, the metric considers the combination of annualised PAYE earnings and raw self-assessed earnings.

³⁵ See www.gov.uk/government/collections/personal-incomes-statistics.

59. Further detail can be found on the Gov.uk website³⁶.

Coverage of the above median earnings or in higher study metric

60. This metric includes all UK-domiciled leavers who are included in the relevant HESA and ILR datasets and have been awarded full Level 4, 5 or 6 qualifications.

61. The metric covers students leaving higher education in academic years:

- 2009-10 (defines the Year 1 year-split metric)
- 2010-11 (defines the Year 2 year-split metric)
- 2011-12 (defines the Year 3 year-split metric).

Exclusions of the above median earnings or in higher study metric

62. The following exclusions apply:

- a. EU and non-EU international students.
- b. Students who were not awarded an undergraduate Level 4, 5 or 6 qualification.
- c. Students who are recorded in another provider's ILR data for the same activity.
- d. Students at providers in Northern Ireland, or at providers that were not required to return HESA or ILR data in 2009-10, 2010-11 or 2011-12.
- e. Students without an HMRC or DWP tax or benefits record in any year, and the small minority of students whose personal characteristics data was not sufficiently complete to enable linking to HMRC or DWP records.
- f. Students who were not in sustained employment or further study three years after graduation.
- g. Students who were in sustained employment three years after graduation but whose salary information was not known.

Benchmarking factors for the above median earnings or in higher study metric

63. The factors used to benchmark the full-time above median earnings or in higher study metric are:

- subject of study
- entry qualifications
- ethnicity
- sex
- disability
- POLAR4 quintile

³⁶ See www.gov.uk/government/statistics/graduate-outcomes-by-degree-subject-and-university.

- level of study.

64. The factors used to benchmark the part-time above median earnings or in higher study metric are:

- subject of study
- age on entry
- entry qualifications
- ethnicity
- sex
- disability
- POLAR4 quintile
- level of study.

Supplementary degree attainment data (at provider level only)

65. Paragraphs 66 to 81 provide a description of the set of supplementary data relating to degree classifications awarded that will be used within the pilot. This data will only be generated at provider level, for providers that hold taught degree awarding powers (TDAPs).

Coverage

66. The supplementary data relating to degree attainment will only be produced for providers that currently hold TDAPs. It covers all awards made by a provider to the students it has taught.

67. The data will include all students who are included in one of the relevant HESA or ILR datasets and recorded as qualifying from a Level 6 (or higher) undergraduate degree award.

68. The data covers UK, other EU and non-EU students.

Exclusions

69. The following exclusions apply:

- a. Students who were not awarded an undergraduate Level 6+ qualification.
- b. Awards made by a provider to students who are registered or taught elsewhere, under franchising or validation arrangements.
- c. Students who are recorded in another provider's ILR data for the same activity.

Benchmarking factors

70. The supplementary data relating to degree attainment is not benchmarked.

Composition of the supplementary data

71. The supplementary data relating to degree attainment will comprise:

- a. Data on differential degree attainment (described further in paragraphs 72 to 75).

- b. Data on grade inflation (described further in paragraphs 76 to 78).
- c. Contextual data on the prior attainment of the graduating cohorts considered by the data on differential degree attainment and grade inflation (described in paragraphs 79 to 81).

Data on differential degree attainment

72. This data will show the proportion of Level 6+ undergraduate degrees awarded as firsts and upper seconds (1sts and 2:1s). The number of students awarded a 1st or 2:1 are counted as the numerator. The denominator counts all Level 6+ undergraduate degree qualifiers awarded a 1st, 2:1, any other degree classification or unclassified degree award.
73. The data will be shown for each of the student groups shown in the standard TEF metric splits (described in Table 5 of the 'Guide to subject-level pilot data'³⁷). For each split, the data will be shown for three cohorts individually: for the cohorts graduating six, four and two years ago. The data covers students awarded qualifications in academic years:
- 2012-13 (defines the 'six years ago' position)
 - 2014-15 (defines the 'four years ago' position)
 - 2016-17 (defines the 'two years ago' position³⁸).
74. The TEF metric workbooks for the pilot will also present the absolute gaps in the proportions between the groups of students in each of these years. Where the differences between the proportions is statistically significant (based on an exact binomial test), these cases will be highlighted within the workbook.
75. This data is at an early stage of development and we expect in future to align it with the Office for Students' wider approach to data on degree attainment, including for the purposes of understanding access and participation in higher education.

Data on grade inflation

76. This data will show the number and proportion of degrees awarded as each of 1sts, 2:1s, other degree classifications and unclassified degree awards 10, six, four and two years ago. The data covers students awarded qualifications in academic years:
- 2008-09 (defines the baseline year, as the '10 years ago' position)
 - 2012-13 (defines the 'six years ago' position)
 - 2014-15 (defines the 'four years ago' position)
 - 2016-17 (defines the 'two years ago' position).
77. Data will be provided for these years, where it is available within providers' HESA and ILR student-level data returns. The Office for Students considers information 'available' for any year in which the provider held and used its own TDAPs: information will only be considered unavailable if the provider did not award any Level 6+ undergraduate degrees to students it taught and registered in a given year. If data is unavailable for the baseline year of 2008-09 but

³⁷ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/.

³⁸ The 'two years ago' position reflects the most recent academic year of HESA and ILR data currently available.

is available for a year between 2008-09 and 2012-13, data will instead be provided for the year that is nearest to 2008-09 as the baseline.

78. The data will also report the percentage changes in the number of graduates awarded each category of degree classification, between 2008-09 (or the baseline year identified) and 2016-17.

Contextual data on the entry qualifications of graduating cohorts

79. The data on differential degree attainment and grade inflation will be accompanied by contextual data on the entry qualifications of students within those graduating cohorts, based on the qualifications those students held at their point of entry to higher education.

80. The data covers students awarded qualifications in academic years:

- 2012-13 (equating to the 'six years ago' position described for the data on differential degree attainment and grade inflation)
- 2014-15 (equating to the 'four years ago' position)
- 2016-17 (equating to the 'two years ago' position).

81. The entry qualification information will present data using the same categories as the TEF contextual data (described in Table 1 of the 'Guide to subject-level pilot data³⁹'). Students are assigned to one of the following categories:

- a. UK-domiciled student holding any higher-education level qualification.
- b. UK-domiciled student with high tariff points (over 390).
- c. UK-domiciled student with medium tariff points (280 to 390).
- d. UK-domiciled student with low tariff points (1 to 280).
- e. UK-domiciled student with non-tariff bearing qualifications.
- f. Non-UK domiciled student. Very few international qualifications are included in the UCAS tariff, so no attempt has been made to sub-categorise the qualifications held by students domiciled outside of the UK.

³⁹ See www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-year-two-2018-19/

Annex C: Additional information for participating providers and TEF student representatives

Timeline

1. Table C1 provides an indicative timeline of activity for providers and Teaching Excellence and Student Framework (TEF) student representatives taking part in the second subject pilot.

Table C1: Timeline for participating providers

Activity	Date	Additional information
Pilot participants receive subject metrics	Week commencing 5 November 2018	Metrics will be posted on the TEF extranet. The OfS expects that metrics will be shared with the TEF student representative and other staff and students involved in writing submissions.
Providers confirm subjects in scope or request changes	From 12 November 2018	See paragraphs 2-7 in this Annex.
Pre-application briefing events	13 November, Birmingham 15 November, London	TEF contacts and TEF student representatives should attend one of these events.
Mid-application feedback event	One day in week of 21-25 January 2019, date tbc	TEF contacts and TEF student representatives should attend.
Early deadline for a sample of submissions	4 February 2019	A subset of pilot participants will be asked for early provider submissions, subject submissions or both (including provider summaries where applicable). These early submissions will be used in panel members' calibration exercises. The providers asked to make early submissions will, in the first instance, be those that participated in last year's pilot.
Submission deadline	25 February 2019	All provider submissions, provider context statements and subject submissions must be uploaded to the TEF extranet in the format specified.
Student declarations	25 February 2019	TEF student representatives will be asked submit a brief declaration to confirm whether students have had appropriate opportunities to be involved in provider submissions.
Post-application feedback event	Date to be confirmed in week commencing 18 March 2019	TEF contacts and TEF student representatives should attend.
Post-application survey and evaluation activity	April-June 2019	Providers should be prepared to take part in feedback and evaluation activities, including a survey, after making their pilot

Providers confidentially receive ratings and statements of findings June 2019

submissions. Details of evaluation activities will be available nearer the time.

Ratings and statements of findings will be uploaded to the TEF extranet.

Requesting changes to data or scope of assessment

2. The OfS will seek to identify in advance and clearly indicate within the TEF metrics workbooks those subjects that are proposed as in and out of scope, as set out in Figure C1.
3. However, it is recognised that the data available to the OfS for this purpose will be neither complete nor sufficiently current to provide a definitive classification of subjects as in or out of scope. During the submission window, providers will have the opportunity to request changes to which of their subjects are in or out of scope. This may include, for example:
 - a. Indicating where subjects the OfS has identified as current and in scope are no longer recruiting new students.
 - b. Confirming whether subjects identified by the OfS as discontinued and no longer recruiting new students have been correctly categorised as out of scope.
4. Any provider that wishes to query the scope classification of any of their subjects will be invited to submit a subject scope declaration form to the OfS, which will then agree any appropriate reclassifications with the provider on a case by case basis. An exemplar declaration form will be published alongside the release of subject-level pilot data⁴⁰.
5. During the submission window, providers may wish to draw the OfS's attention to errors in their underlying data and begin the process for requesting amendments to this. However, providers should note that the OfS will not be able to process amended data or reissue subject-level pilot metrics to providers during the subject pilot submission window.
6. Outside the subject pilot submission process, a request to amend data for TEF purposes can be made in the usual way⁴¹.
7. The process for requesting changes to data or scope of assessment for the subject-level pilot is presented at Figure C2.

⁴⁰ This document will become available at www.officeforstudents.org.uk/advice-and-guidance/teaching/subject-level-pilot-2018-19/further-technical-guidance in early November 2018.

⁴¹ See www.officeforstudents.org.uk/data-and-analysis/amendments-to-data/.

Figure C1: Which of my subjects will be assessed in the second subject pilot?

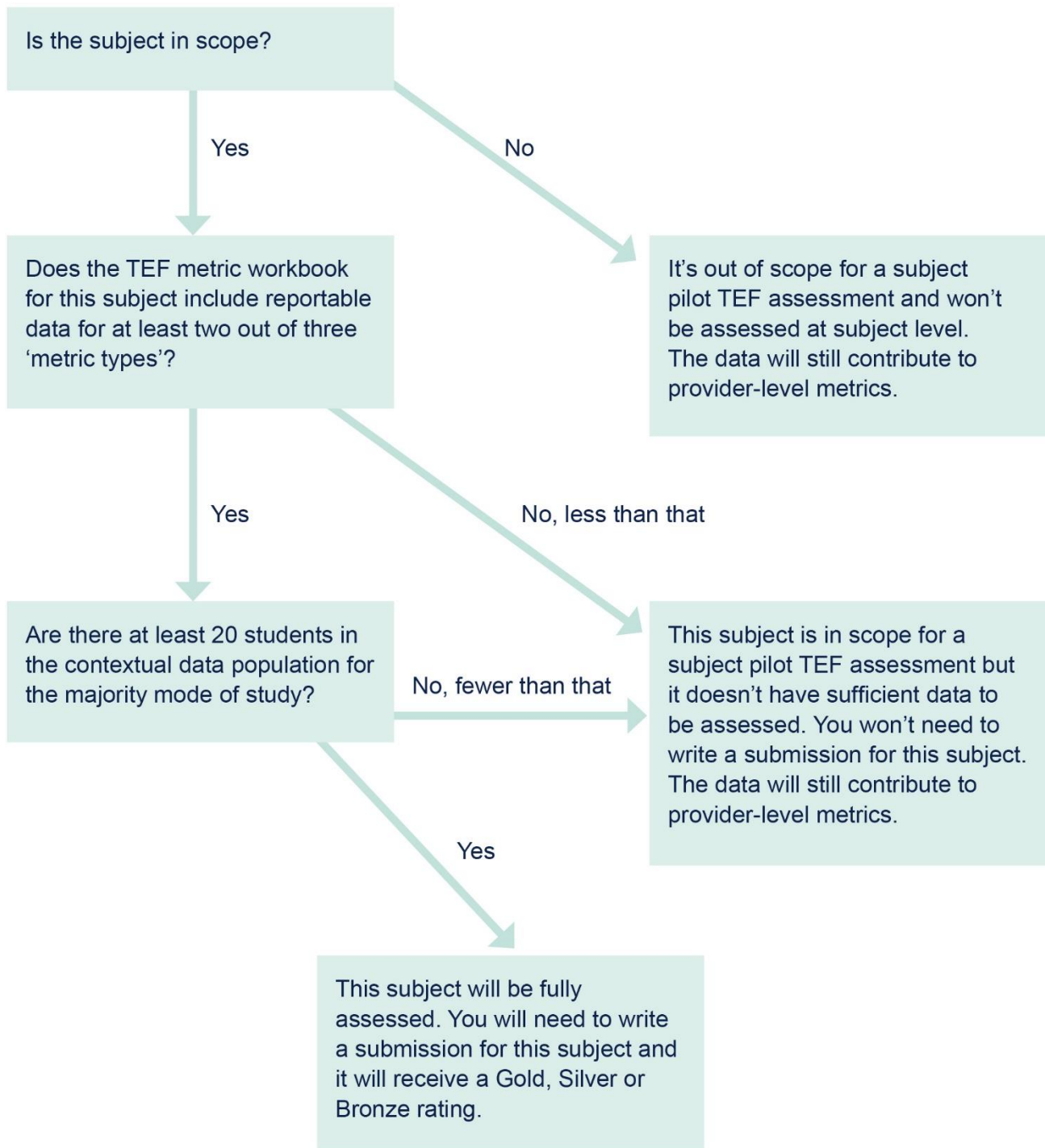
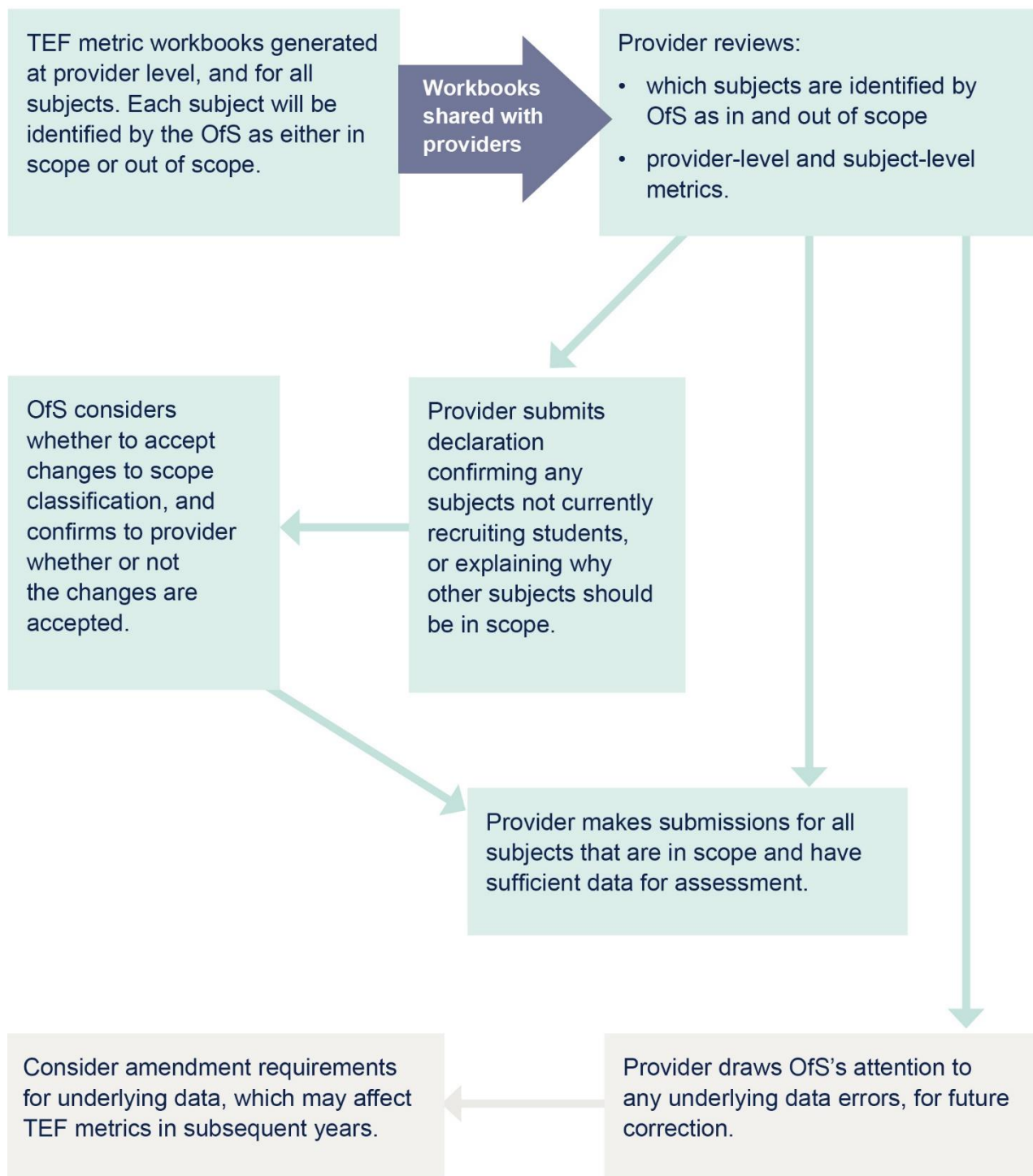


Figure C2: Process for reviewing which subjects are in scope



Submission instructions

What to submit and format requirements

8. Participating providers will be expected to submit:
 - a. A provider submission of not more than 15 pages.
 - b. A subject submission of not more than five pages, for each subject that is being assessed.

- c. A single provider summary statement of not more than two pages.
9. Providers with at least 35 per cent of students studying part-time may optionally submit an additional page of quantitative information about part-time provision, at provider level only. It will be made clear in the TEF metrics workbook whether or not a provider is eligible to submit this additional page.
10. All provider and subject submissions, provider summary statements and additional part-time pages must adhere to the following stylistic and formatting conventions:
- Arial font 11 point (minimum)
 - 2cm page margins (minimum)
 - single spacing (minimum)
 - provider name and UK Provider Reference Number in a header on every page
 - page numbers in the footer.
11. Templates for each type of document that adhere to these formatting requirements will be provided by the OfS.
12. Footnotes are permissible, either to clarify statements in the documents or to indicate where the submitted evidence has been drawn from. However, hyperlinks to primary evidence should not be included, and the panel members will be instructed not to access any referenced sources or follow any hyperlinks in a submission. Judgements will be based only on the information as set out in Table 10, and no additional external evidence. The onus therefore is on the provider to ensure that all the information required to make the judgement (in addition to the metrics and data supplied by the OfS) is included in the submissions, provider summary statement or additional page of part-time data, where relevant.
13. No appendices or any other type of information may be included if not incorporated within the page limits.
14. All documents must be submitted in PDF format. As they have to be accessible to screen-reading technology, scanned PDFs are not acceptable.

When to submit

15. Some providers will be asked to make early submissions for part of their provision. This could be either the provider submission or provider summary statement and a small selection of subject submissions. This is to make available example submissions in good time for the panel member calibration exercise. Providers will receive as much notice as possible of early submission requirements. To reduce burden, providers asked to make early submissions will as far as possible be those that participated in the first year of pilots.
16. All submissions, provider summary statements and additional part time pages must be uploaded to the subject pilot area of the TEF extranet by noon on the following dates:
- **Monday 4 February 2019** for early submissions

- **Monday 25 February 2019** for all other submissions.

How to submit: TEF extranet

17. Providers need to use the TEF extranet to access their metrics workbooks and upload their subject pilot submissions. The TEF extranet is available at <https://tef.officeforstudents.org.uk/submissions>.
18. Each provider has been informed of the unique group keys that are needed to access the TEF extranet user groups. A hard copy letter, dated 03 October 2018, was sent to the head of institution or accountable officer with this information. The head of institution or accountable officer is responsible for disseminating these group keys to the appropriate individuals, taking care to ensure that they have the responsibility and authority to access the content provided via the TEF extranet. Pilot participants will receive an additional group key giving access to the pilot upload site after their participation in the pilot is confirmed.
19. Queries about access to the TEF extranet should be sent to tefmetrics@officeforstudents.org.uk.

Submitting the student declaration

20. The TEF student representative at each participating provider will be invited to submit a brief declaration to the OfS, to confirm whether student representatives have had appropriate opportunities to be involved in the pilot submissions. A template for this will be developed in discussion with TEF student representatives, and provided along with instructions for how to submit the declaration, by January 2019.
21. TEF student representatives will be asked to submit their declarations by **Monday 25 February 2019**.

Clarification and verification procedure

22. Panel members will, by exception, be able to request clarification or verification of critical information already contained in a provider or subject submission.
23. Panel members may raise queries only where the response could potentially make a material difference to the overall outcome.
24. It is the provider's responsibility to include in its submission all the information required for panel members to make a judgement. Consequently, queries will not be raised – and responses will not be considered – that seek to expand on or add new evidence to a submission.
25. The decision to request clarification or verification from the provider will be made at the discretion of the head of TEF or the TEF subject pilot manager. Queries will normally be directed during March to April 2019 to providers' TEF contacts, who should be available to answer queries if necessary.
26. The clarification and verification process will be kept under review throughout the second pilot, with a view to establishing whether it is necessary to include this process in a future subject-level TEF exercise.

Annex D: Subject panel roles and key responsibilities

Table D1: Main Panel roles

Role	Key responsibilities
Chair	<ul style="list-style-type: none"> To oversee the work of the Main Panel and facilitate and report on feedback from across the panels. To chair meetings ensuring collective decisions on provider-level ratings are reached, and that advice and guidance to the subject panels is agreed.
Deputy chair (academic)	<ul style="list-style-type: none"> In addition to acting as a main panel member, to assist and deputise for the chair, and observe subject panel meetings.
Deputy chair (student)	<ul style="list-style-type: none"> In addition to acting as a main panel member, to assist and deputise for the chair, and observe subject panel meetings. To collate and report on feedback from student panel members.
Members (academic and student)	<ul style="list-style-type: none"> To review a caseload of provider-level metrics and submissions. To act as lead reviewer for some of these, and draft their statements of findings. To contribute to meetings of the Main Panel including taking collective decisions on provider level ratings; considering the consistency of subject-level assessments; and providing feedback on the pilot process.
Widening participation (WP) experts	<ul style="list-style-type: none"> To review a sample of provider-level metrics and submissions to identify issues related to WP. To provide expert input and advice on widening participation to the Main Panel and the WP liaisons on the subject panels. To contribute to meetings of the Main Panel including taking collective decisions on provider level ratings; considering the consistency of subject-level assessments; and providing feedback on the pilot process.
Employment experts	<ul style="list-style-type: none"> To review a sample of provider-level metrics and submissions to identify issues related to employment. To provide expert input and advice on employment-related issues to the Main Panel. To contribute to meetings of the Main Panel including taking collective decisions on provider level ratings; considering the consistency of subject-level assessments; and providing feedback on the pilot process.
Subject panel co-chairs and deputies	<ul style="list-style-type: none"> To provide a link between the Main Panel and subject panels, reporting to the Main Panel on progress and findings from the subject panels. To contribute to meetings of the Main Panel including taking collective decisions on provider level ratings; considering the consistency of subject-level assessments; and providing feedback on the pilot process.

Table D2: Subject panel roles

Role	Key responsibilities
Co- chairs (academic)	<ul style="list-style-type: none"> • To oversee the work of the subject panel, including allocating assessment material to panel members. • To chair meetings ensuring all panel members' contributions are valued and collective decisions are reached. • To check for consistent application of standards across the subject panel. • To review subject-level metrics and submissions for the purposes of moderation and resolving borderline cases. • To convey advice and guidance from the Main Panel, supporting consistency of approach across the subject panels. • To take an overview of the lessons learned and report feedback on the process.
Deputy chairs (student)	<ul style="list-style-type: none"> • To assist the co-chairs in overseeing the work of the subject panel, checking for consistency and conveying guidance from the Main Panel. • To review subject-level metrics and submissions for the purposes of moderation and resolving borderline cases. • To deputise by chairing parts of meetings. • To contribute to reporting on lessons learned, and lead on reporting any student-specific feedback.
Members (academic and student)	<ul style="list-style-type: none"> • To review a caseload of subject-level metrics and submissions. • To act as lead reviewer for some of these, and draft their statements of findings. • To contribute to meetings of the subject panel, including taking collective decisions on subject-level ratings, and providing feedback on the pilot process.
Employer and PSRB representatives	<ul style="list-style-type: none"> • To review a sample of subject-level metrics and submissions, contributing in particular input and advice on employer and professional perspectives. • To contribute to meetings of the subject panel, including taking collective decisions on subject-level ratings, and providing feedback on the pilot process.
WP liaisons	<p>These will be academic or student members of the subject panel with the following additional responsibilities:</p> <ul style="list-style-type: none"> • To direct the subject panel's attention to, and advise on, WP issues. • To liaise with the WP experts on the Main Panel and each other, to support consistent approaches to WP issues across the subject panels and provide additional feedback on WP issues.
Interdisciplinary liaisons	<p>These will be academic members of the subject panel with the following additional responsibilities:</p> <ul style="list-style-type: none"> • To direct the panel's attention to, and advise, on interdisciplinary provision. • To liaise with each other to share expertise and experience across the subject panels, and provide additional feedback on the assessment of interdisciplinary provision.



© The Office for Students copyright 2018

This publication is available under the Open Government Licence 3.0.

www.nationalarchives.gov.uk/doc/open-government-licence/version/3/