

Consultation Response

OfS Consultation on regulating quality and standards in higher education

Submitted: 25th January 2020

Consultation: <https://www.officeforstudents.org.uk/publications/consultation-on-regulating-quality-and-standards-in-higher-education/>

Proposal 1: Define 'quality' and 'standards' more clearly for the purpose of setting the minimum baseline requirements for all providers

Question 1a: Do you agree or disagree with the proposed definitions of 'quality' and 'standards' set out in Table 1 of Annex A (see above) and that this should be used to express minimum baseline requirements for quality and standards in revised B conditions?

Agree/Disagree/Neither agree nor disagree

Comment:

Table 1 is not a definition of 'quality' or 'standards'; it is a collection of measures, indicators, desirable outcomes and expressions of principle.

Using a simplistic and superficial conception of what constitutes quality and standards in HE could drive behaviour which is counterintuitive to improving quality and standards. A focus on the student experience and enhancement initiatives may be lost if definitions of what constitutes 'good quality' become overly reliant on disaggregated data. It is possible that a provider could meet the baseline requirements, including as to outcomes, without in fact providing a high quality educational experience.

It is not clear how numerical baselines or indeed other data can be identified to evidence many of the aspects included in Table 1. Any data indicators that do underpin Table 1 will fluctuate across time for a range of reasons, and those changes do not in isolation indicate there is threat to quality or standards. In our view whether an institution is meeting the "standards" set out in the statements in Table 1 can only be assessed by reviewing the steps that an institution is taking to meet them, i.e. by reviewing the process that the institution has taken to ensure it is meeting them.

We do not agree that measures of outcomes should be included in Table 1. Outcomes are dependent on many factors, linked to background, opportunity, geography, prior attainment and context. We do not believe that there is a sufficient causal link between quality in higher education and these outcomes to justify their inclusion in a definition of quality or in the regulatory framework. For example, there is not necessarily any link between physical and virtual learning resources and high continuation rates. Conversely, good continuation rates do not necessarily indicate a robust, high quality education experience. In this context, we note the content of the recent policy note issued by the Higher Education Policy Institute (<https://www.hepi.ac.uk/2021/01/07/a-short-guide-to-non-continuation-in-uk-universities/>) ."*putting undue weight on the issue may end up as an example of 'Goodhart's Law', which is commonly expressed as 'When a measure becomes a target, it ceases to be a good measure.'*". We suggest the same argument applies to the other proposed outcomes measures.

We also do not agree that, if quantitative measures are used, they should be absolute numbers. Measures used should be benchmarked. We note this from the Pearce Review;

- *"To make fair judgements about an institution's excellence, student intake characteristics should be taken into account (i.e. data should be benchmarked) but panels should also have clear information on absolute performance to inform the overall assessment. In doing this however, if there is a conflict between absolute and benchmarked performance, **benchmarking should continue to take precedence over absolute values**".*

We note that for the TEF, the Pearce Review recommended that location of employment should be used as an additional benchmarking factor for graduate outcomes because of the regional variation in employment outcomes and we believe that this is the case if such data is used in the quality framework.

There are also other metrics set out in Table 1, for example the number of staff involved in delivery. There are no proposals for how this will be assessed. We are concerned that this will become a separate metric in its own right.

We share the concerns raised by UUK in relation to the use of degree classification descriptors (<https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Pages/response-office-for-students-consultation-regulating-quality-and-standards.aspx>).

As is noted in paragraph 46 of the consultation document, references to the Quality Code will be removed from the guidance in the regulatory framework and the Quality Code is not referred to in Table 1. The consultation suggests that this will remove complexity and lack of clarity in the OfS requirements, however, it raises questions about the purpose, role and status of the Quality Code going forwards. We suggest that the OfS need to clarify whether and what the ongoing role of the Quality Code is if these changes are made. The Quality Code is cross UK and this supports the strength of the UK HE brand internationally. Removal of references to the Quality Code from the regulatory framework would lead to significant differences in how quality is assessed between England and the other devolved nations. As noted above, whether an institution is meeting the “standards” set out in the statements in Table 1 can only be assessed by reviewing the steps that an institution is taking to meet them, i.e. by reviewing the process that the institution has taken to ensure it is meeting them, which are currently supported by the Quality Code and its supporting guidance. This will be replaced by increasingly prescriptive process requirements from the OfS – we note for example the approach taken in the pandemic in which the OfS have set out very specific process instructions such as to review our compliance with consumer protection requirements, and present them to the board. In our view this level of prescriptive process instruction is not consistent with a principles based regulatory approach that claims to be based on outcomes rather than prescriptive process.

Question 1b: Do you have any comments about how the proposed definitions of quality and standards set out in Table 1 of Annex A should be assessed for individual providers?

Comment: No further comment. Please see our response to 1a

Question 1c: Do you agree or disagree with the proposal in paragraphs 41 to 43 to express initial requirements differently from the equivalent ongoing requirement for providers seeking registration?

Agree/Disagree/ Neither agree nor disagree

Comment:

If, as we suggest, outcomes are not part of the definition of quality, this problem does not arise.

We suggest that if the proposal to include such data is adopted, the lack of such data for some suppliers should be reflected in ongoing monitoring, i.e. closer scrutiny of other indicators, and use of internal data to provide earlier indications of performance given lag in published data. In our view whether an institution is meeting the “standards” set out in the statements in Table 1 can only be assessed by reviewing the steps that an institution is taking to meet them, i.e. by reviewing the process that the institution has taken to ensure it is meeting them.

Proposal 2: Set numerical baselines for student outcomes and assess a provider’s absolute performance in relation to these

Question 2a: Do you agree or disagree with the proposed approach to assessing student outcomes set out in Annex B?

Agree/Disagree/ Neither agree nor disagree

Comment:

We do not agree that numerical indicators for student outcomes should be included. The three proposed indicators are neither appropriate nor sufficient evidence of whether providers are meeting the other baseline requirements set out in Table 1.

We note that the proposal in Annex B is to use a 5 year time series and, if numerical indicators are used (which we do not agree with) then we support the use of 5 years of data. If indicators are used then a longer time

frame, not just a snapshot, is important. The longer term view and the trajectory are important indicators of quality, which may be obscured by something that impacts a particular indicator in a particular year. This is especially the case if the data is not benchmarked, because there is a greater risk of volatility in outcomes when measured against an absolute number that is not contextualised. We note that, for example, for the TEF, the Pearce Review recommended that location of employment should be used as a benchmarking factor for graduate outcomes because of the regional variation in employment outcomes.

It is not clear what will be used as the definition for 'Completion', further detail is required on how this is going to be measured to be able to comment meaningfully, particularly if the intention is to move from projected outcomes as currently performed by HESA PIs to use actual outcomes, which would introduce a significant time lag into the data.

We are concerned about the process by which the numerical baselines will be set, and changed, over time. We note that in Annex B it is stated that the OfS will be setting "higher and more challenging baselines". We are concerned that the setting of such baselines will be driven by a political, rather than a regulatory process and will not be informed by the sector. We are concerned that, as with the TEF, the baselines will be set using a normal distribution in such a way that will deliberately ensure that a proportion of the sector will "fail", particularly if there is no benchmarking.

Question 2b: Are there any other quantitative measures of student outcomes that we should consider in addition to continuation, completion and progression (see Annex B paragraph 18)?

Comment:

Although we do not agree that quantitative measures of student outcomes should be included in the framework, if they are, we suggest that consideration should be given to a 'value added' measure – entry qualification vs degree outcome. We note that in degree classifications analysis, the OfS have chosen to describe outcomes that are not linked to prior attainment as "unexplained" – we believe that these outcomes are not unexplained, they are the result of the hard work put in by staff to improve quality and learning outcomes for students and the hard work of students engaging positively with the opportunities they are given and therefore this is an important indicator of quality. It has proved difficult to measure learning gain (although further consideration of this was recommended by Dame Shirley Pearce in her review of the TEF), but "value add" could be a good proxy.

Question 2c: Do you agree or disagree with the proposals for the levels of study at which indicators should be constructed? Should any additional indicators be considered (see Annex B paragraph 25)?

Agree/Disagree/ Neither agree nor disagree

Comment:

Although we do not agree that quantitative measures of student outcomes should be included in the framework, if they are:

- there should be a minimum threshold for sample sizes for data validity as some subject areas are relatively small, and there should be a threshold for inclusion according to the proportion of that provision in the portfolio as a whole.
- it may be useful to split out Foundation programmes and apprenticeship provision as this provision is quite different and distinct, although these may well relate to small groups.

We note that the Pearce Review of the TEF, and the ONS report that it commissioned, are highly critical of the implications of using small data sets in the TEF, and we argue that the same concerns apply in the quality framework.

We are concerned about the use of subject groups as provision does not map clearly or cleanly onto programmes and the data will therefore therefore lack transparency and meaning. When considering subject groups for the TEF it was evident that the subject definitions bring together programmes from across different faculties and departments which are delivered and supported in quite different ways. Therefore it is difficult and potentially misleading to draw conclusions on quality from subject metrics, it will be burdensome for providers to review this data and difficult to draw meaningful conclusions or develop actions from it. We

believe that this has been recognised by the government in their decision to drop plans for subject level TEF and we urge the OfS to reconsider plans to review indicators at subject level.

Question 2d: Do you have any comments about an appropriate balance between the volume and complexity of indicators and a method that allows us to identify ‘pockets’ of performance that are below a numerical baseline (see Annex B paragraph 32)?

Comment:

Although we do not agree that quantitative measures of student outcomes should be included in the framework, if they are, it is very important that the data doesn’t become too complex and that any reported metrics are transparent enough to provide a measure of the provision in question.

Subject level definitions do not map clearly or cleanly onto programmes and subject level data will therefore lack transparency and meaning. When considering subject groups for the TEF it was evident that the subject definitions bring together programmes from across different faculties and departments which are delivered and supported in quite different ways. Therefore it will be difficult and potentially misleading to draw conclusions on quality from subject metrics, it will be burdensome for providers to review this data and difficult to draw meaningful conclusions or develop actions from it. We believe that this has been recognised by the government in their decision to drop plans for subject level TEF and we urge the OfS to reconsider plans to review indicators at subject level.

In addition there needs to be a set minimum threshold for sample sizes so that data used is meaningful.

The government has made it clear that it is expecting a focus on reducing the burden of bureaucracy across the sector, and for example, the government response to the Pearce Review specifically states that there should be an additional principle of proportionality and that the costs of the new TEF “*should also be kept proportionate and for each exercise the costs, for both providers or the OfS, should, at an absolute maximum, not exceed the costs per provider of the TEF exercise that has taken place to date*”. We are concerned that, in contradiction to this principle the burden for providers in responding to the data and other requirements proposed by this framework will be substantially higher than has been the case before.

Question 2e: Do you agree or disagree with the demographic characteristics we propose to use (see Annex B paragraph 36)? Are there further demographic characteristics which we should consider including in the list of ‘split indicators’?

Agree/Disagree/ Neither agree nor disagree

Comment:

Although we do not agree that quantitative measures of student outcomes should be included in the framework, if they are, the demographic characteristics used must be aligned to the APP dataset and definitions to aid simplicity and transparency and ensure additional burden is not introduced by requiring institutions to understand and monitor another set of similar and overlapping metrics.

Again, minimum sample size thresholds for data to be reported should apply to split data.

Question 2f: Do you agree or disagree that the longitudinal educational outcomes dataset should be used to provide further indicators in relation to graduate outcomes (see Annex B paragraph 46)?

Agree/Disagree/ Neither agree nor disagree

Comment:

We have said that we do not believe that numerical outcomes should be included in the framework and we do not support the use of subject level data for the reasons outlined above. The time lag for all such outcomes data and in particular LEO makes it difficult to read across into current student experience or institutional performance

We note that for the TEF, the Pearce Review recommended that location of employment should be used as a benchmarking factor for graduate outcomes because of the regional variation in employment outcomes and we believe that this is the case if such data is used in the quality framework.

If numerical indicators are included, including at subject level we would strongly support the use of LEO, as LEO is a more robust source of data than the Graduate Outcomes data, particularly for earnings data, even though it has a time lag, and in that case we also support the use of a five year data series to provide context for this data.

Question 2g: Do you have any comments about how the range of sector-level performance should be taken into account in setting numerical baselines (see Annex B paragraph 57)?

Comment:

We are concerned about the process by which the numerical baselines will be set, and changed, over time. We note that in Annex B it is stated that the OfS will be setting “higher and more challenging baselines”. We are concerned that the setting of such baselines will be driven by a political, rather than a regulatory process and will not be informed by the sector. We are concerned that, as with the TEF, the baselines will be set using a normal distribution in such a way that will deliberately ensure that a proportion of the sector will “fail”, particularly if there is no benchmarking.

We note that the new proposals for the TEF now include a fourth category of providers failing to show sufficient evidence of excellence. We are very concerned about this, as it suggests that there will be a substantial overlap between the TEF and the quality framework although the same data will be used for both. The TEF was more informative when it was applied only to those providers who met the quality baseline rather than constituting an alternative quality framework in its own right – using much of the same data but with a different time frame and a different assessment process. The new structure will be very confusing.

Not benchmarking is problematic. By not benchmarking, the framework is ignoring factors that are outside of the institution’s control, e.g. the wide range of societal, regional and economic factors that influence these outcomes. We note that, for example, for the TEF, the Pearce Review recommended that location of employment should be used as a benchmarking factor for graduate outcomes because of the regional variation in employment outcomes.

We note that although it was not mentioned in the Government response to the Pearce Review of the TEF, the review itself recommended the continued use of benchmarking in the TEF, and we believe that the same arguments apply to the quality framework.

Not benchmarking will also mean that pockets of excellence in terms of relative outcomes are ignored and that there is no opportunity to identify and share good practice.

Question 2h: Do you have any comments about the other contextual factors that should be taken into account and the weight that should be placed on them (see Annex B paragraph 68)?

Comment:

The consultation document does not set out how or in what circumstances this contextual information would be gathered. As described this is a significant set of additional data and reporting requirements which are not reflected in current reporting or the separate reportable events consultation – for example (c) and (d) - evidence of action taken and future plans. Providing this data would be a significant additional burden on providers.

If this information is required, the OfS should review their current transparency reporting and other data requirements to ensure that as much of the necessary data as possible is collected in a consistent way on a regular annual cycle with sufficient advance notice. There will be a significant regulatory burden for providers if additional data is requested on an adhoc basis at short notice to respond to requests that are triggered by particular events.

We are concerned about potential requirements for data which overlaps with, or is collated in a different way for other pre-existing reporting requirements, requirements for data which we do not routinely collect, or that is not collected in a consistent way by different providers.

There is a risk that institutions will be monitoring and reporting on multiple overlapping data sets for the quality framework, the TEF and the APP, which will increase the regulatory burden and make it harder for providers to interpret and apply conclusions from the data.

Proposal 3: Clarify the indicators and approach used for risk-based monitoring of quality and standards

Question 3: Do you agree or disagree with the proposals in Annex C for monitoring ongoing compliance with regulatory requirements for quality and standards?

Agree/Disagree/ Neither agree nor disagree

Comment:

The range of evidence that is described in the consultation is a significant set of additional data and reporting requirements which are not reflected in current reporting or the separate reportable events consultation. The consultation document does not set out how or in what circumstances this information would be gathered and while the consultation states that this data is available to the OfS we do not believe that all of it is (eg complaints, external reviews). In practice, as we have highlighted above, we believe that a lot of the evidence that the OfS will require will, in practice, relate to evidence of processes, and yet this is in the context of a steadily diminishing set of guidance about process and supporting frameworks for providers.

If this information is required, the OfS should review their current transparency reporting and other data requirements to ensure that as much of the necessary data as possible is collected in a consistent way on a regular annual cycle with sufficient advance notice. There will be a significant regulatory burden for providers if additional data is requested on an adhoc basis at short notice to respond to requests that are triggered by particular events.

We are concerned about potential requirements for data which overlaps with, or is collated in a different way for other pre-existing reporting requirements, requirements for data which we do not routinely collect, or that is not collected in a consistent way by different providers (such as informal complaints).

The proposals are unclear about the relationship between the standards and the TEF. We are concerned that there will be overlap, and inconsistency, between this data and the data required for the TEF and it should be clear how the TEF builds on the baseline set out in the standards. We note that the new proposals for the TEF now include a fourth category of providers failing to show sufficient evidence of excellence. We are very concerned about this, as it suggests that there will be a substantial overlap between the TEF and the quality framework although the same data will be used for both. The TEF was more informative when it was applied only to those providers who met the quality baseline rather than constituting an alternative quality framework in its own right – using much of the same data but with a different time frame and a different assessment process. The new structure will be very confusing. We are concerned about the proposal to undertake further investigation in the case of a TEF award in the lowest category, given that the use of the normal distribution in the TEF flagging process forced a number of providers into the lowest category.

There is a risk that institutions will be monitoring and reporting on multiple overlapping data sets for the quality framework, the TEF and the APP, which will increase the regulatory burden and make it harder for providers to interpret and apply conclusions from the data.

In our view whether an institution is meeting the “standards” set out in the statements in Table 1 can only be assessed by reviewing the steps that an institution is taking to meet them, i.e. by reviewing the process that the institution has taken to ensure it is meeting them. This undermines the suggestion that the framework will be based on outcomes data – it won't. Collating and providing that information will represent a substantial investment of time by institutions.

We are concerned by the suggestions about reportable events in this consultation which are more granular and inconsistent with the separate consultation on reportable events. The OfS should consider the additional regulatory burden of any extension to the list of reportable events. For providers that take their responsibilities to report very seriously, as we do, each individual new category of reportable event does add a significant regulatory burden. To ensure a provider is meeting its duty this requires considerable institutional effort from training, identification, collation, reporting for consideration, advising, Accountable Officer decision-making to Board reporting and involvement. We therefore believe it is proportionate to reserve the use of reportable events to those matters that are genuinely material, urgent and of institutional level significance. Reportable Events should not be used as a data gathering exercise for more routine or operational issues because of the

burden this places on all levels of an institution given the very compressed timescales to identify, evaluate, determine and report. This is not proportionate to the potential likelihood of harm to students.

The government has made it clear that it is expecting a focus on reducing the burden of bureaucracy across the sector, and for example, the government response to the Pearce Review specifically states that there should be an additional principle of proportionality and that the costs of the new TEF “*should also be kept proportionate and for each exercise the costs, for both providers or the OfS, should, at an absolute maximum, not exceed the costs per provider of the TEF exercise that has taken place to date*”. We are concerned that, in contradiction to this principle the burden for providers in responding to the data and other requirements proposed by this framework will be substantially higher than has been the case before.

Proposal 4: Clarify our approach to intervention and our approach to gathering further information about concerns about quality and standards

Question 4: Do you agree or disagree with the proposals in paragraphs 86 to 101 for our approach to intervention and gathering further information about concerns about quality and standards?

Agree/Disagree/ Neither agree nor disagree

Comment:

As we have noted above, as described in the consultation this is a significant set of additional data and reporting requirements which are not reflected in current reporting or the separate reportable events consultation. The consultation document does not set out how or in what circumstances much of this information would be gathered.

If this information is required, the OfS should review their current transparency reporting and other data requirements to ensure that as much of the necessary data as possible is collected in a consistent way on a regular annual cycle with sufficient advance notice. There will be a significant regulatory burden for providers if additional data is requested on an adhoc basis at short notice to respond to requests that are triggered by particular events. The government has made it clear that it is expecting a focus on reducing the burden of bureaucracy across the sector, and for example, the government response to the Pearce Review specifically states that there should be an additional principle of proportionality. We are concerned that, in contradiction to this principle the burden for providers in responding to the data and other requirements proposed by this framework will be substantially higher than has been the case before. We are concerned about potential requirements for data which overlaps with, or is collated in a different way for other pre-existing reporting requirements, requirements for data which we do not routinely collect, or that is not collected in a consistent way by different providers (such as informal complaints).

The proposals are unclear about the relationship between the standards and the TEF. We are concerned that there will be overlap, and inconsistency, between this data and the data required for the TEF and it should be clear how the TEF builds on the baseline set out in the standards. There is a risk that institutions will be monitoring and reporting on multiple overlapping data sets for the quality framework, the TEF and the APP, which will increase the regulatory burden and make it harder for providers to interpret and apply conclusions from the data.

We are concerned by the suggestions about reportable events in this consultation which are more granular and inconsistent with the separate consultation on reportable events.

All Proposals:

Question 5: Do you have any comments about any unintended consequences of these proposals, for example for particular types of provider or course or for any particular types of student?

Comment:

We believe that the result of these proposals will be to drive behaviours that are not about quality and standards, i.e. distilling quality down to single numerical measure will focus providers on achieving a narrow set of numerical outcomes at the cost of a wider focus on a holistic definition of quality.

Furthermore, this plus the lack of benchmarking would create a disincentive to providers to recruit under-represented groups due to the impact this may have on institutional baselines.

An institution may still add substantial value locally despite not consistently achieving sector wide numerical baseline standards.

We do not believe that this approach sits well alongside the APP action plans to reduce gaps. The APP reflects the longer term approach that is required to address gaps - providers have put in plans over an initial 5 year period to take action to close these gaps. The proposed framework does not recognise or incentivise a longer term approach to these issues.

Question 6: Do you have any comments about the potential impact of these proposals on individuals on the basis of their protected characteristics?

Comment:

We believe that the result of these proposals will be to drive behaviours that are not about quality and standards, i.e. distilling quality down to single numerical measure will focus providers on achieving a narrow set of numerical outcomes at the cost of a wider focus on a holistic definition of quality.

Furthermore, this plus the lack of benchmarking would create a disincentive to providers to recruit under-represented groups due to the impact this may have on institutional baselines.

An institution may still add substantial value locally despite not consistently achieving sector wide numerical baseline standards.

We do not believe that this approach sits well alongside the APP action plans to reduce gaps. The APP reflects the longer term approach that is required to address gaps - providers have put in plans over an initial 5 year period to take action to close these gaps. The proposed framework does not recognise or incentivise a longer term approach to these issues.

Question 7: Do you have any comments about where regulatory burden could be reduced?

Comment:

As noted above we are concerned by the regulatory burden that will be created by these new arrangements and the additional information that will be required. Data gathered on a planned annual cycle for providers with appropriate advance notice also creates less regulatory burden than repeated urgent short notice requirements.

Question 8: Do you have any other comments?

Comment:

N/A

JANE FORSTER

VC's Policy Advisor

Follow: [@PolicyBU](#) on Twitter

|

SARAH CARTER

Policy & Public Affairs Officer

policy@bournemouth.ac.uk

|