

Russell Group response to the consultation on the Future Research Assessment Programme (FRAP)

Summary of key points and recommendations

1. The primary purposes of the REF should be recognising and rewarding excellence in research, determining the allocation of QR and equivalents in the devolved administrations, to provide the evidence base and accountability for public investment in research, and benchmarking.
2. Research culture and environment are critical for research excellence, but we recognise there are challenges in measuring these objectively. REF is an important, but not the sole driver of a positive research culture. The Environment statement is the appropriate place to evidence progress on research culture issues.
3. We recommend a simplified research Environment statement comprising an institution-level template using contextualised metrics, and a reduced-length unit-level statement for local activities and measures of success.
4. The REF could take steps to recognise the value of team-based research and a wider range of people who contribute to research in the unit-level environment statement
5. The REF should retain the balance of assessing the components of excellence as: Outputs (60%), Impact (25%), Environment (15%).
6. The REF should retain the 7-year REF timetable to facilitate long-term, strategic planning. We do not support the introduction of an interim, metrics-based assessment, or a panel-staggered rolling exercise.
7. Guidance for the REF rules should be published well in advance of the REF exercise
8. For the assessment of outputs, panels should be encouraged to maintain optional use of bibliometric data to inform peer review in certain UOAs where these are well validated
9. The current metrics used in assessing impact are sufficient - peer review should remain the primary method of assessment
10. We recommend a review of the effectiveness of flagging interdisciplinary research and consider integrating improved mechanisms to recognise and reward it
11. We recommend a removal of self-declaration and special circumstances and the introduction of a “tolerance band” mechanism for zero submission
12. We propose that a pilot study is run to assess whether outputs could be sampled for panels to review, using REF2021 data to test the sensitivity of dipstick sampling method on UOA scores

1. The purpose and role of the REF

The primary purposes of the REF should be recognising and rewarding excellence in research, determining the allocation of QR and equivalents in the devolved administrations, to provide the evidence base and accountability for public investment in research, and benchmarking.

- 1.1 The Research Excellence Framework (REF) is a fundamental part of the UK's dual support system for research funding. The dual support system, as protected in the Higher Education and Research Act 2017, guarantees an unhypothecated stream of quality-related research funding (QR), and equivalents in the devolved administrations, which allows universities to engage in long-term strategic planning and to respond quickly to emerging opportunities. The combination of low-bureaucracy, agile funding and competitively awarded grants provided through the dual support system ensures the diversity and breadth of research in the UK and maintaining it in the next REF is crucial to continue to underpin the UK research base.
- 1.2 Lord Stern's independent report on the Research Excellence Framework shows that over the years it has been implemented, research assessment exercises in the UK (RAE/REF) have been successful in driving a sustained improvement in the quality of research. The next REF exercise should not try to fulfil multiple purposes but instead retain a clear focus on maintaining this quality and excellence in research.
- 1.3 The primary purposes of REF should be:
 - **Recognising and rewarding excellence:** as an external assessment which seeks to recognise and reward the very highest levels of excellence in research through robust peer-review assessment.
 - **Allocation of QR and equivalents in the devolved administrations:** as the mechanism to determine the allocation of QR funding.
 - **Accountability for investment:** providing an evidence base and accountability for public investment in research.
 - **Benchmarking:** as providing a peer-reviewed benchmark information that can be used in a variety of ways, including at an international level.
- 1.4 Internally, REF results can be used to inform decisions on resource allocation and future planning within the institution where QR flows to departments in whole or in part based on REF results, and in supporting department-level student recruitment, especially at the PGR level. The exercise can also be used to share models of best practice around issues such as EDI and Open Research through the unit level Environment statements. There may be scope to make greater use of the information collected via REF to support the wider development and application of best practice, in universities and across the research eco-system.
- 1.5 REF, with peer review at its heart, is a robust way of allocating QR funding and its equivalents in the devolved nations; this function cannot be fulfilled by other routes. The robustness of the exercise is also key for providing accountability for investment to government, particularly given the long-term nature of research. For this reason, by validating the unhypothecated nature of QR, REF is critical to the dual support system which gives the UK the edge against international competitors. This directly supports the Government's ambitions around becoming a Science Superpower and Innovation Nation.

2. Research culture and environment in the REF

Research culture and environment are critical for research excellence, but we recognise there are challenges in measuring these objectively. REF is an important, but not the sole driver of a positive research culture. The Environment statement is the appropriate place to evidence progress on research culture issues.

- 2.1 A positive research culture and environment is critical to support research excellence, but we recognise that there are challenges in measuring these objectively. The REF undoubtedly has an impact on driving behaviour in universities and has an important role to play in creating a positive research culture.
- 2.2 At the same time, REF is and should not be the sole driver of a positive research culture. There are already mechanisms in the wider R&D system which support research culture, including the requirements in the concordats, research grant terms and conditions, and the publication of EDI data and other indicators. The REF should seek to align with and complement these other drivers, not seek to replicate, or work against them.
- 2.3 We note that the REF already measures some aspects of the research environment that form part of research culture, for example open access and open data, but there is scope to go further in a new type of Environment statement, as we propose below.

Reforms to the Environment section to capture research culture

We recommend a simplified research Environment statement comprising an institution-level template using contextualised metrics, and a reduced-length unit-level statement for local activities and measures of success.

- 2.4 We propose that the Environment statement is the appropriate place to capture and evidence positive research culture and environment. However, in discussion with members we have found that identifying robust metrics or indicators for research culture is challenging, as research culture does not lend itself well to metrics. Here we propose some themes and suggestions of possible metrics but suggest that ultimately the REF leaves it at institutions' discretion how they use this to persuade the reader that a positive culture is being developed. Metrics alone can never be sufficient to assess complex research environments, and therefore should always be accompanied by contextual information.
- 2.5 Whatever assessment criteria are chosen to illustrate research culture, they must allow review panels to discriminate across different quality standards. We recognise that with any metric there is the potential for gaming, and therefore recommend careful consideration to ensure they are meaningful and responsible. We must also be mindful of suggesting metrics to assess Environment beyond what we already collect as this risks unintended consequences on researcher behaviour.

An institution-level environment template

- 2.6 The purpose of the institution-level environment template would be to serve as an assurance of compliance, with metrics or indicators to demonstrate the institution has the necessary policies and procedures to support a healthy research environment and culture, as well as the mission and disciplinary mix of the institution and its research strategy. It should have a clear remit and guidance about what to include, and we propose that a good starting point for elements relating to research culture would be to use existing sector concordats and agreements.

- 2.7 The template could include some suggestions of metrics or indicators, with the opportunity to provide some qualitative narrative information alongside. Themes that might be included are:
- **Researcher careers:** illustrating how universities support and develop research careers and skills, e.g. PGR completion rates within tuition-fee paying period, the proportion of externally-funded research staff on open-ended contracts, development for ECRs and processes for inducting staff, CPD opportunities and level of engagement, and how the university supports mobility within and outside the university.
 - **Research integrity:** demonstrating the university's environment for integrity in research, for example, action plans in response to the concordat for research integrity, the identity of the named person for research integrity, the policy for handling concerns, and an annual integrity assurance report which evidences progress.
 - **Team-based research:** demonstrating how universities are recognising and rewarding the contributions of a wide range of people to research, for example, the adoption of the CReDIT taxonomy (or similar); evidence of diverse and inclusionary practices in teams
 - **Open research:** for example, evidence of the uptake of open research practices
 - **EDI:** indicators of the EDI profile of the workforce and the students via HESA data, measuring the change over time.
 - **Staff surveys:** for example, action taken and progress made against issues raised in staff surveys
- 2.8 We recognise that research culture is challenging to measure through metrics alone, and so a narrative option for institutions remains the most appropriate way to present information for assessment. HEIs could be given the flexibility to add additional metrics of their own on top of these, wherever possible selecting those that are in the public domain to reduce burden and increase consistency.
- 2.9 The Environment section comprises more than Research Culture issues and People alone, e.g. research strategy, research facilities, opportunities for collaboration and environment, including research income and research degrees awarded. These sections allow for a more objective, data-driven assessment. Consideration could be given to allowing the output and impact scores to provide baseline environment scores which are then adjusted up or down by evidence of environment attributes.

A short unit-level statement

- 2.10 The local environment is key to supporting meaningful improvements in research culture, and a single institutional statement would be too general to reflect the different areas of strength in disciplines. We propose pared-down appendixes to the institutional template at the unit level which would serve to reflect the local environment and allow for the specifics of that discipline and how they have contributed to, and benefited from, their institutional environment.
- 2.11 Unit-level statements should be short-form response style, supported by evidence in the form of metrics or indicators, rather than the current extended prose format. This will reduce burden and allow assessment to focus on the content of the information provided. Consideration could be given to including different questions for different main panels, since the environments are necessarily different.

The REF could take steps to recognise the value of team-based research and a wider range of people who contribute to research in the unit-level environment statement

- 2.12 One way in which the next REF could support a positive research culture is to take steps to recognise the value of a wider range of people who contribute to research, encouraging a more inclusive research community and recognising that not all disciplines work in teams in the traditional sense. Visible adoption of the CReDIT taxonomy (or similar) would be an indicator, as well as diverse and inclusionary practices in teams. Metrics could be introduced around the number of authors and co-authorship networks. This could be aggregated at the UOA level, without allocating an output to them.
- 2.13 We had considered the possibility of the next REF broadening eligibility of FTE to incorporate a wider range of people, however the resultant potential for complexity, burden and gameplay means that we would not recommend this at this stage. We had also considered whether introducing a broader selection of outputs might go further to support inclusive and team-based research, by moving away from the published paper as the dominant type of output and taking advantage of the platforms created through open research / open science. However, before exploring this idea further we propose that the guidance for the next REF seeks to encourage the full use of the 21 output types already permitted (and the 'other' category) and evaluate whether there is evidence to support the addition of new types of output.

3. Weighting the components of excellence

The REF should retain the balance of assessing the components of excellence as: Outputs (60%), Impact (25%), Environment (15%).

- 3.1 We support retaining the current balance of assessing the components of excellence as: Outputs (60%), Impact (25%), Environment (15%) in the next REF exercise. There is strong consensus amongst members for this position.
- 3.2 The Outputs section is the most objective and robust element of REF, and therefore should remain the dominant part of the scoring to fulfil REF's primary purpose as a driver of quality-based funding underpinning the dual support system. Any consideration of changing the weighting of the components of excellence in future exercises should be supported by evaluation and backed with evidence.

4. Frequency and sequencing

The REF should retain the 7-year REF timetable to facilitate long-term, strategic planning. We do not support the introduction of an interim, metrics-based assessment, or a panel-staggered rolling exercise.

- 4.1 The REF should retain its ability to facilitate long-term, strategic planning, which aligns with the strategic intent of QR, the Russell Group therefore supports retaining a 7-year cycle. In funding terms, this length of cycle provides sustainability and resilience for the sector and recognises the nature of discovery research as a long-term endeavour for which the delivery of impact is almost always longer than a typical political cycle. Maintaining a 7-year term also ensures consistency with previous exercises.
- 4.2 We strongly recommend prioritising the long-term stability offered by a 7-year exercise over the currency of information which a more frequent exercise might offer, as this puts trust in the institution and enables strategic prioritisation. We do not support arguments for a more frequent exercise to maintain currency of information as the hypothecated nature of QR allows universities the freedom to respond to emerging priorities alongside long-term

investment decisions. In addition, other types of funding including grant funding are able to support and reward rapidly strengthening disciplines, allowing QR funding to focus on long-term institutional support.

- 4.3 Furthermore, we do not support any suggestion to hold an annual process, nor do we support any interim, metrics-based assessment processes at subject level. Introduction of this type of evaluation risks encouraging short-term thinking and metrics-chasing over research excellence which would adversely impact the quality of research in the UK. There is also a risk that assessments on this basis may start out as light-touch and evolve into something more burdensome which is of little benefit.
- 4.4 Russell Group members have considered whether a panel-staggered process could be implemented, whereby main panels could work to different cycles. While there were some views in support of this type of sequencing, overall, we believe the cons significantly outweigh the pros. For example, splitting the exercise by main panel risks leading to disciplinary silos, stifling cross-disciplinary research centres, increasing the burden on departments which submit to different main panels as they will always be in the process of submitting to REF, as well as implications for staggering QR allocation which would adversely affect the stability of funding and ability to strategically invest in priority areas over time, hampering the ability of QR to support institutional-level strategic thinking. For these reasons, we would not support a panel-staggered REF.

5. Publication of guidance

Guidance for the REF rules should be published well in advance of the REF exercise

- 5.1 We strongly recommend that guidance for the REF rules be published well in advance of the exercise to allow universities time to prepare. There should be fewer adjustments or pilot exercises during the REF cycle itself, and any pilot exercises should be completed well in advance of the next exercise. HEIs will need sufficient time to understand and implement the new guidance, and late notice of the rules has been a source of burden in past exercises. In future, if the REF guidance were to be published in advance of the assessment period rather than at the end of it, this would give HEIs time to begin to collect any necessary data.
- 5.2 Concerns have been raised that the FRAP consultation process, with further rounds of consultation, may mean universities are a long way through the assessment period before the final guidance is ready. If the consultation process means early release of guidance will be challenging, depending on the degree of change made to the next exercise, we propose two ways to mitigate this:
 - **Extension of the assessment period:** A one-year extension to the assessment period could be implemented to allow universities more time to adapt to new guidance. With additional time, there should be less need for any updates to the guidance and institutions would have longer to prepare for the next REF in line with any new rules.
 - **Delaying significant changes to after the next REF:** Depending on the degree of changes, consideration might be given to implementing *some* significant changes at the next REF and delaying any other significant changes until the exercise after that. This applies to any possible metrics not readily available to avoid any retrospective data collection which is ineffective and inefficient.

6. Metrics

For the assessment of outputs, panels should be encouraged to maintain optional use of bibliometric data to inform peer review in certain UOAs where these are well validated

- 6.1 REF is already very effective at assessing **outputs** and we do not see a strong argument for extending the use of metrics beyond what was permitted within REF2021. The element of fair judgement against the specific REF criteria that peer review enables is essential when evaluating a diverse range of outputs across the research spectrum – and even more so when considering the quality of interdisciplinary research. As such, we would propose that sub-panels should again be able to decide whether they would find it useful to have bibliometric data made available to them. We would strongly recommend that any metric data supplied to sub-panels must only be used to inform peer review (as in REF2021).
- 6.2 We propose that improvements could be made at the next REF on transparency and calibration. For example, when one single output is submitted by multiple institutions it would be a useful calibration to know that it has received the same score in all the submissions to which it was attached; use of DoI tagging should make this relatively easy. Then the system could report back on the number of outputs submitted by more than one institution and the score allocated (preserving anonymity of output).

The current metrics used in assessing impact are sufficient - peer review should remain the primary method of assessment

- 6.3 The current metrics used in assessing **impact** are sufficient; peer review of individual case studies must remain the primary method of assessment here. The diversity of impacts delivered as a result of excellent research makes comparisons and judgements very difficult unless by peer review. Updated guidance for REF2021 on the standardisation of metrics within impact case studies, and the collection of funding information within the additional contextual data, has been a welcome step to introduce consistency.
- 6.4 Impact should not be driven by a solely quantitative, metrics-driven assessment but should also draw on qualitative assessments which can better capture the role and value of partnership and engagement beyond academia and the significance of more nuanced outcomes. Summative impact metrics can be inferred from KEF and HEBCIS so duplication of this assessment should be avoided in REF.

7. UOAs and interdisciplinary research

We recommend a review of the effectiveness of flagging interdisciplinary research and consider integrating improved mechanisms to recognise and reward it

- 7.1 At the next REF, impact and outputs should continue to be assessed at the level of the 34 expert-level UOAs based on disciplinary areas under the four main panels. We would not support the use of self-defined research themes, and we maintain that disciplinary areas are the most useful way of categorising research.
- 7.2 REF should not value interdisciplinary research above disciplinary excellence; the objective should be to ensure that all work is equitably assessed. Russell Group members have discussed the issues faced by interdisciplinary research and feedback has suggested that the REF2021 mechanism of a yes/no flag for interdisciplinary research outputs did not work well and could be cumbersome. Leaving the definition of interdisciplinary research to the discretion of institutions meant that it was interpreted very differently by institutions. Therefore, a review of its effectiveness and exploration of improved mechanisms would be welcome.

- 7.3 One improvement which could be explored for the next REF is introducing a **systematic flagging of research outputs**, whereby every output is categorised or flagged as either ‘within sub panel’, ‘within main panel’, or ‘across main panels’. This approach would have the advantage of working within the existing disciplinary sub-panel approach, without the need for new interdisciplinary panels or cross-referral between panels.
- 7.4 We would not support any suggestion of the introduction of thematic panels into the REF. There are several drawbacks with this approach, including ensuring the outputs are assessed by the relevant experts and inconsistency of decision-making on which panel to submit to.

8. Burden and bureaucracy

- 8.1 As a mechanism to underpin the distribution of QR funding, it should be recognised that the REF is already a relatively cost-effective exercise. The cost of REF 2014 has been estimated to be roughly 2.4% of the £10.2 billion in QR research funds projected to be distributed between 2015-16 and 2020-21¹, although we recognise the costs are higher when compared to the change in funding distribution between exercises. Whilst the cost of the exercise to institutions in terms of staff time is high at key points in the REF cycle, we would say that the cost and burden is proportionate when taken in the overall context of the return on investments and efforts from QR, when compared with preparing bids for research grants.²
- 8.2 While there is scope for burden and bureaucracy to be reduced further, we must emphasise that any progress here retains the full assessment of outputs to maintain the REF’s credibility; we would not support cost-reduction as a reason to move away from peer review as the primary assessment method. Below we set out ideas which will serve to reduce burden on institutions whilst maintaining the robustness of the exercise:

We recommend a removal of self-declaration and special circumstances and the introduction of a “tolerance band” mechanism for zero submission

- 8.3 Declaration of special circumstances presents a significant burden for universities, too great a focus on outputs for individuals and a misguided assumption that HEIs expect individual staff to produce outputs at a similar rate without due regard for personal circumstances. For this reason, and as part of emphasising the collective rather than individual achievement, we propose that all self-declaration and special circumstances are removed. This is in the interests of collective assessment and a positive culture around REF as well as a reduction of burden.
- 8.4 An alternative mechanism to consider might be the introduction of a **“tolerance band” mechanism for zero submission** whereby universities are given a tolerance band for the number of permitted zero submissions, with a certain amount of leeway permitted as a percentage of submitted FTE. This would allow institutions to use existing HR processes to manage special circumstances internally, and removes the burden associated with collecting and providing this evidence to the REF.

¹ REF accountability review: costs, benefit and burden. Technopolis for HEFCE (July 2015).

² University A estimates that it cost £7M in staff time to prepare its REF return, which led to receiving around £350M in QR from 2014-2021. The university estimates that it spends about £10 M annually in staff time to win £120M of competitively won research funding annually, so in simple terms the ratio of bidding cost to funding awarded, is 1:12 for individual grant funding and 1:50 for QR, so REF generating QR is more than four times as cost-efficient.

We propose that a pilot study is run to assess whether outputs could be sampled for panels to review, using REF2021 data to test the sensitivity of dipstick sampling method on UOA scores

- 8.5 It has been suggested that the peer review process for REF2021 was highly burdensome on panel members. Member feedback has been that more can be done by REF to ensure that future panel members are aware of the time commitment required in advance.
- 8.6 As an idea to address the level of burden for panels, a pilot study could be conducted to test whether panels could, in future, only peer review a proportion of submitted outputs. This pilot could assess whether this approach would lead to significantly different results compared to the current approach of assessing all outputs, and whether an algorithm could be designed to implement sampling in a fair way.
- 8.7 A dipstick sampling method, whereby a university puts forward a pool of outputs which are randomly selected from for peer review could significantly reduce the burden on panel members, especially if coupled with an expansion of panel membership. To test the sensitivity of this method, an experiment could be run at different levels amongst different disciplines from the pool of outputs using REF2021 data across all panels, and the variance this produces for the UOA's overall scores. If the volume of outputs to be assessed could be at least halved, this approach could be worth introducing in future exercises.
- 8.8 There are challenges with this method which would require further investigation. For example, this approach may lend itself better to larger panels, and there could be a challenge in this methodology for disciplines where monographs/non-paper-based outputs are the norm. Examination of any unintended consequences of this method should therefore be undertaken before approved for implementation.

Other suggestions for reducing burden and bureaucracy:

- **Removal of the requirement for income in-kind data:** Member feedback has been that income in-kind data is not systematically collected by either HEIs or UKRI, and thus is not robust or auditable. It plays only a minor role in the assessment of the environment for only a few units, but is burdensome to prepare and validate for all, and so removing this requirement would be a welcome step.
 - **Clear guidance in a single web portal:** Another suggestion for reducing burden is to have guidance that is less verbose and complex, without the overlap between multiple documents and web pages (e.g., for REF 2021 there were c.20 guidance documents in addition to web-based information). This can be achieved by having a single web portal hosting the guidance, which can be added to or updated as necessary. The guidance should be written in consultation with experienced REF panel secretariat.
 - **Open Access:** REF 2021 played an important role in driving progress towards an open access culture, but member feedback has been that it did so through a particularly bureaucratic process that was labour-intensive to implement and monitor. For many researchers, this has resulted in open access being considered a bureaucratic chore rather than something to be valued. OA can be better incentivised through a simpler and less bureaucratic approach, to be as closely aligned to the UKRI Open Access Policy as possible.
- 8.9 Members had considered other ways to reduce the burden on universities associated with the selection of outputs and in order to encourage a cross-institution assessment. However,

although this idea seemed to have merit in principle, we estimated it would be too technically challenging to implement and had potential for adverse unintended consequences.

May 2022