



Universities UK

Policy Briefing

The future of research
assessment: the principles of
reform

Policy Briefings

This new series of Policy Briefings published by Universities UK will provide authoritative and accessible analyses of current and emerging higher education policy issues. We aim to publish at least six booklets a year on major topics of the day, with an analysis of an issue, identification of policy options and, where relevant, the Universities UK or sector position. The booklets will draw on existing Universities UK policy work as well as new research that has been undertaken or commissioned.

The future of research assessment: the principles of reform

Executive summary

- This briefing is a contribution to the debate about the future of research assessment that began in March 2006 when the Government outlined proposals for the assessment and funding of research using 'metrics' – quantitative indicators – rather than peer judgements.
- The need to build on Britain's exceptional research performance, which is based on the foundation of the dual support system, should be the starting point of any discussion of new research assessment mechanisms.
- Fundamental reform of the Research Assessment Exercise (RAE) is now needed.
- The RAE has placed significant pressures on individuals and institutions and often distorts academic and institutional culture and activity.
- The RAE has produced diminishing returns over time.
- Planning for the RAE in 2008 is well advanced and the decision to proceed as planned was the right one.
- It is essential that the outcomes of the next RAE are the primary evidence of quality, informing funding for at least a four-year period.
- There are major concerns about the exclusive use of income metrics as proposed by the Government and other quantitative measures need to be used.
- Peer involvement should remain within a new assessment process although metrics will have a key role to play.
- Discussion of the way forward should be based on a set of agreed principles and a Universities UK statement (Annex A) has been agreed as basis for developing new assessment arrangements.
- Four possible outline approaches to replacing the RAE are offered for debate (Annex B).

1 Introduction

1.1 The Chancellor of the Exchequer's announcement in his March 2006 budget speech that the Research Assessment Exercise (RAE) would be ended has sparked an intense debate about its replacement. The Government outlined proposals for the assessment and funding of research mainly using 'metrics' – quantitative indicators – rather than peer judgements. Its plans were further developed in a consultation document *Reform of higher education research assessment and funding* published by the Department for Education and Skills (DfES) in June 2006.

1.2 This policy briefing summarises the views of Universities UK on the Government's proposals, explains the principles on which change should be based and suggests some possible ways forward. This paper complements its formal submission to Government which focuses on the specific questions raised in the consultation document.

1.3 There is no doubt that fundamental reform of the RAE process is now needed and that many of the Government's concerns about the problems associated with the RAE are shared by the higher education sector. There is, however, a need to broaden the scope of the debate about research assessment and funding, and a statement of key principles that should shape further policy discussion in this area has been put forward by Universities UK. These principles, which have informed this briefing, are listed in Annex A and cover the following dimensions:

- Quality Research (QR) funding;
- quality assessment and peer involvement;
- universality;
- applied and user focused research;
- behavioural impact;
- transition and financial stability;
- cost and simplicity; and
- policy fit.

2 UK research performance

2.1 Any discussion about the future of research assessment needs to be set in the context of Britain's exceptional research performance. The UK maintains its place as the second most important producer of high quality research in the world. The UK produces nine per cent of the world's scientific papers with a citation share of 12 per cent, second only to the USA, and has continued to increase its share of the world's most influential papers, from 12.9 per cent to 13.2 per cent¹. The continued strength of UK research is evident across the full range of scientific disciplines from engineering and physical sciences to the arts and humanities.

2.2 British universities have been the foundation of this success, which has been achieved with lower investment than their international competitors, despite increased public funding in recent years. UK research is not only strong internationally but is also relevant to research users at regional and national levels. Any new mechanism should seek to sustain and build upon this outstanding success.

2.3 The success of the UK's university research has been underpinned by the dual support system, which provides public funds to institutions in two streams – one as part of their block grant provided by the higher education funding councils, and the other in the form of project-based grants, provided by research councils. A key strength of the system is that the funding council grant is unhypothecated allowing university leaders the freedom to take strategic decisions about the focus of their research effort. It also means that there are multiple sources of funding for research, with multiple decision points about the research that should be supported and where resources should be concentrated. This creates a healthy and dynamic research base. The dual support system is the essential foundation of UK research performance and Universities UK welcomes the Government's commitment to maintain it as the basis of a reformed system.

3 The need for reform

3.1 Under the current arrangements QR funds are allocated by the funding councils on the basis of quality judgements through the periodic RAE. The RAE is an *ex post* system, based on informed peer review. A study comparing systems for university research evaluation and funding, conducted by the Science Policy Research Unit (SPRU)² in 2003, shows that the UK has developed one of the most advanced research evaluation systems in Europe. It is one of only a few countries that has implemented a performance-based system to allocate research funding, although others are moving in this direction. The study highlights the clear advantages of this approach. These have been evident in the RAE, which since its introduction in 1986 has:

- provided a firm basis for the selective funding of research, based on excellence;
- created a strong incentive to improve individual as well as institutional research performance;
- encouraged the development of institutional management and strategic planning processes and increased efficiency;
- provided greater accountability for public funds invested in research (as opposed to a system based, for example, on student numbers and institutional size).

3.2 However, in recent years there has been a growing dissatisfaction with the RAE and increased calls for reform both from the sector and in Whitehall. The two key drivers for the reform of the RAE, as expressed in the Government's *Next Steps* document, have been the burden on the sector and the distortions it has created.

3.3 Before considering these criticisms it should be recognised that it is difficult to distinguish between general dissatisfaction with the RAE and the way in which the outcome has been used selectively to allocate funds and the impact this has had on individuals, teams and institutions. The results of the RAE2001 exceeded expectations, but the failure of the Government, particularly in England, to fund fully the results caused extreme concern to the academic community. Some 64 per cent of research submitted in 2001 was found to be of national or international excellence (compared with 43 per cent in the previous RAE) but this was not matched by increased funding. These changes have had a significant impact on the finances of those institutions affected by the cuts. It remains unclear how the outcomes of the next RAE in 2008 will be used to distribute funding.

3.4 The direct costs of RAE2001 (ie, costs to the funding councils) came to £5.6 million, and the expected figure for RAE2008 is £10-12 million³. Calculating the indirect costs to universities is more problematic. There are difficulties in identifying the specific costs of RAE submissions as opposed to the continuing costs of managing research activities in a university. A survey of the costs of the RAE in 1996 produced an estimate of indirect costs of £30 million. A later study produced an estimate (including opportunity costs) of £37.5 million for higher education institutions (HEIs) in England – or 0.8 per cent of the total funds allocated on the basis of the RAE results⁴. The Higher Education Funding Council for England (HEFCE) believes that the costs to HEIs of RAE2008 will not be radically different. The results of the next exercise will be used to allocate some £8 billion of research funding. If the total cost of the assessment were £45 million, this would represent around 0.6 per cent of the resources allocated, comparing favorably with the costs associated with project-based grant allocations using expert review through the research councils.

3.5 However, it is important not to overlook the substantial ‘wear and tear’ caused by RAE, particularly in the pressures that it places on individuals and institutions. The problems this has caused are not insignificant. The five-year RAE cycle has come to dominate and often distort academic and institutional culture and activity. The distorting effects of the RAE are well documented and include⁵:

- pressure on universities to employ staff who can immediately contribute to the RAE;
- bias against applied research;
- difficulties in assessing applied and inter-disciplinary research;
- danger of grade inflation;
- incentives for games playing producing distortions in institutional decision-making;
- separation of research from teaching, creating an academic culture where teaching has a lower status;
- barriers to new approaches, with ‘safe’ research often being rewarded;
- lack of accountability between RAEs; and
- use of researchers as panel members reducing their research output.

3.6 A SPRU study⁶ of the costs and benefits of the RAE concluded that it had produced diminishing returns over time. Initially, due to the investment needed to set up and adapt to a new system, the costs of a performance-based approach will outweigh the benefits it brings. In time the benefits will grow as universities develop clearer goals and strategies, and efficiency and performance are increased. After a number of iterations, however, the returns will diminish as strategies are in place and poor performance has been addressed. Running further exercises will mean that benefits fall away at an ever-increasing rate.

3.7 The SPRU report suggests that for the RAE the benefit curve has peaked and diminishing returns have set in. The scope for further gains by introducing additional processes is limited. Whilst the SPRU study looks primarily at the direct costs outlined in paragraph 3.4, it is also possible to categorise the distorting effects outlined above as additional costs, which would become more difficult to justify as the benefits diminish. This assessment provides a compelling argument for reform, but one that indicates that any new system should maintain and build upon the benefits the RAE has realised.

3.8 Universities UK therefore endorses the Government’s conclusion that fundamental reform of the RAE is now required. Reducing the burden of the research assessment process on the sector would be an essential consequence of change. The effort involved in transforming the system would, of course, need to be proportionate to the gains. The need to lighten the burden will also need to be weighed against the requirement for a robust system that is fit for purpose and has the confidence of the academic community. Cost reduction is not the only driver for change, and the current review provides a crucial opportunity to examine how the problems associated with the RAE can be addressed.

4 Government proposals

The next RAE

4.1 Although Universities UK supports the need for fundamental reform, the planning for RAE2008 is well advanced and the decision to proceed as planned was the right one under the circumstances. It will be important to resolve quickly the issue (as raised in the Government's consultation paper) of whether the panels should make greater use of metrics in 2008. Panel criteria and working methods have already been published and institutions' planning is well advanced. There is a strong case for minimum disruption to RAE2008.

4.2 For this reason there are considerable difficulties with the Government's proposal that any new system should be phased in from 2009. It is essential that the outcomes of RAE2008 are the main evidence of quality, informing funding allocations for at least a four-year period. This will ensure that there is relative stability in the system, and that confidence in the coming exercise is maintained. It will also provide sufficient time to develop and test any new system in parallel without causing disruption. It is important to recognise that a new system, which attempts to address the criticisms of the RAE, particularly in relation to applied and user-focused research, is likely to produce different funding outcomes. Therefore, once a new system is agreed, is fit for purpose and has the confidence of the academic community its introduction must be managed and moderated to avoid any destabilising effects.

Metrics

4.3 Universities UK supports the selective funding of research based on quality. It is essential that some element of peer involvement should remain within a new assessment process although metrics will have a key role to play. Most of the metrics suggested to date are proxy indicators of quality, rather than measures of quality. Peer involvement should therefore continue, for example, in overseeing the validity and selection of the metrics and in interpreting and moderating their outcomes. This would not, however, mean reinventing the current peer review system, with its myriad of panels.

4.4 The main advantages of an increased focus on quantitative indicators are that they would introduce greater simplicity and transparency as well as reducing the bureaucratic burden and cost. The major concerns that need to be addressed include the fact that metrics may not adequately cover all disciplines (for example the arts and humanities and the social sciences) and could lead to unforeseen behavioural changes and distortions. This would suggest that the extent to which metrics can be relied upon may vary between disciplines. Another key challenge will be to identify and develop the metrics themselves. Bearing these points in mind the relevant peer communities should, at a minimum, be involved in selecting appropriate metrics across broad discipline areas.

The main indicators

4.5 The following paragraphs outline thoughts on some of the key indicators that could be included in a new mechanism. Good indicators emerge naturally from the research process, so they must be linked to one or more of its stages: inputs > activity > outputs > outcome. When developing indicators it is also important that the following considerations are taken into account:

- relevance;
- the feasibility of data collection;
- the quality of the data; and
- the meaning of any metrics generated from the data – in particular, the values that indicate improvement or excellence.

The ability to obtain useful data is another important consideration. The Higher Education Statistics Agency (HESA) collects data from institutions to a specified format and covers funding, students and staff. Output data are collected from a number of sources that include HESA (postgraduate research students) and others that do not involve institutions (such as book catalogues, journal articles and patents). The most robust metrics relate back to peer review, which most academic researchers identify as the 'gold standard'.

4.6 The five proposals put forward by the Government are based on income metrics, which are input measures, and therefore represent five variations on a single model. There are major concerns about the exclusive use of income-based metrics but the rest of this section considers the merits of the other indicators that might be used.

4.7 Esteem indicators are problematic: the data are difficult to identify and often impossible to validate. They would be unreliable as foreground measures, or when used in isolation. They are also difficult to scale: no one claims to have been invited to give a 'minor' presentation at a conference, nor do they give a record of 'fleeting visits' by overseas researchers. Nobel Prizes are clearly highly competitive and peer reviewed but not many people receive them. Journal editorships are an option, as are senior roles in learned societies. However, there are eminent people who have done none of these things. Prizes from learned societies could be a good indicator but only because they are normally not measured. If they became a national indicator then they would proliferate and it would then be necessary to make difficult judgements as to which were genuine and significant. Activity measures are less informative in themselves, although the presence of a substantial postgraduate research population is usually a reflection of a beneficial research environment.

4.8 Output and outcome measures have a more useful role to play. PhD awards are a good output measure because of the examination process but it would be useful to have improved information about subsequent employment. Publications and their impact tend to be the focus for metrics. They have the advantage of always having the key attributes of time, location and discipline so they feed readily into statistical comparisons. Patents are a specialised form of output, but are difficult to analyse and the information is not universally linked to origin.

4.9 A marker of 'performance' is citation analysis of journal articles. This can provide good coverage, soundly based data and a link to quality. There is, however, a possibility that changes to the structure of publication through 'open access' will in due course supersede traditional journals. Developments in this area will bring their own advantages and create opportunities for the development of new indicators. Much of the work in this area, as well as the opportunities it presents, is not currently on the radar of policy makers. It is essential that this is explored further before decisions are made. A key question is the extent to which citation metrics are appropriate as a resource allocation model because of the impact that this could have on behaviour. As with input measures, it is likely that they could be used productively as part of a basket of indicators in a supporting role.

Limitations of income-based metrics

4.10 As stated above, the Government's five proposals are based on income metrics. There are a number of objections to this approach. Crucially, a mechanism that directly relates all funding council research support to the income earned through grants and contracts can no longer be regarded as dual support and would not be consistent with the Government's commitment to maintaining the system.

4.11 A justification for a greater reliance in income metrics was provided in the Government's *Next Steps* document. It was suggested that an analysis of research income at institutional level shows a high degree of correlation between grant and contract income and QR income, thus justifying the need to end the parallel peer review process that takes place in each part of the dual support system. Although the analysis works at institutional level, it does not apply so well at subject level. It also assumes a high degree of homogeneity across the sector and fails to recognise differences in institutional size. There is also concern about the impact on the arts and humanities, although such concerns could apply to some other subject areas.

4.12 A key problem is the fact that the proposals fail to link funding to quality. The consultation document recognises that the quality measures generated by the models are not very highly correlated with individual RAE ratings. Our own modelling confirms this⁷. Although there may be concerns about using the RAE as a 'benchmark' for a new system it would still call into question the

extent to which the models can claim to generate true 'quality indicators'. A more robust link to the quality of research than is evident in the Government's proposals is required.

4.13 This would mean developing a more sophisticated way of reflecting the value of applied research to ensure that it is properly recognised and supported. The reliance on income as an indicator could provide some misleading results in view of the fact that some types of research are more costly than others and that some collaborative work may be undertaken on a low cost basis. Indicators of quality as well as quantity that can recognise outputs will be needed. There will be challenges in developing robust output and impact measures, but it will be important to work with the community (both academic and user) to understand how these can be developed and incorporated into the system.

4.14 Under an income-based model there would also be a real danger of driving up volume. A key question when using income as a metric is the impact that it would have on grant applications to the research councils and the real possibility of over-subscription that it raises. Although the move to full economic costing will reduce this risk as institutions look to manage their portfolios on a more sustainable basis it should not be underestimated. Income from the research councils and other funders also fluctuate. An advantage of QR based on the periodic RAE is that it has allowed for a stable planning horizon.

4.15 An income-based system would also fail to recognise the contribution of those without explicit financial support. This would work against the appointment and progression of less established 'early stage' researchers. There is also concern that each of the five models in the consultation would create an unsustainable transfer market for those academics with a particularly successful track record of winning grants.

4.16 It is essential that any new system applies the same assessment framework to all research activity while recognising subject distinctiveness. Different mechanisms that produce different quality measures for some discipline areas or even for specific types of institution could prove divisive and difficult to operate. It will be hard to draw any meaningful 'line' between subject areas and it could create barriers to collaboration and have a negative impact on multi-disciplinary research. It could also lead to an ossification of the system if different parts of the research base become 'compartmentalised'. Any system has to provide scope for change and evolution in the research base.

5 The way forward?

5.1 The decision to proceed with RAE2008 and the need to ensure that the outcome informs funding for a period of up to four years provide ample time to ensure that a new assessment system is designed and fully assessed before it is adopted. Although the RAE needs to be reformed it has been an important factor in enhancing university research performance over the past 20 years and in enabling the UK to maintain its relative international standing in research. It is vital that the successor arrangements continue to support and enhance university research productivity.

5.2 Universities UK believes that future discussion of the way forward should be based on a set of agreed principles and it has endorsed the statement in Annex A as a basis for developing new assessment arrangements. It is too early in the debate to offer a specific alternative to the RAE as much more work is needed, but four possible outline approaches can be readily identified. They are summarised in the chart at Annex B. There is a range of possibilities extending from a pure metrics-based assessment at one end of the spectrum to a streamlined peer review system that has major input from metrics at the other. The intervening options assume a (varying) degree of peer involvement.

5.3 Two models would be worth exploring further. One approach (Option C) is based on the use of 'trigger' metrics, which could be employed either as part of a rolling review process or to determine inclusion in, or exclusion from, a fixed cycle review. This model would use RAE2008 as the baseline, with a further assessment only taking place if metrics show a significant change outside a set tolerance band. Funding might stay in steady state until the metrics indicate that a fuller assessment or audit was needed. Appropriate trigger metrics could be selected and monitored by a panel type structure. Such peer involvement would need to be significantly scaled down and standardised in comparison to the current RAE process (ie, it could be based on the 15 main panels, while abandoning the sub panels). These panels would also undertake any fuller review, if and when it is required.

5.4 This option is not without its difficulties. Certainly for the rolling approach it might be necessary to maintain panels 'at the ready' on a permanent basis, which would not help to reduce bureaucracy in any significant way. Submissions for a trigger-based review would need to be made on a similar basis to the RAE and there would be no reduction in the burden for those institutions. Also both trigger approaches could lead to games playing with institutions trying to manipulate the system for the benefit of league table standing ie, to get an assessment triggered in order to move up the table. This would not meet the criteria of developing a less deleterious approach in terms of behavioural distortions. These problems might be addressed if the panels established that there was a prima facie case for an assessment and ensured that metrics would only trigger an assessment when significant change occurred.

However, another key problem would be that the outcomes would not provide a reliable picture across the board, as the assessment process could become stratified. This lack of a 'level playing field' could also be open to legal challenge.

5.5 Another possibility (Option B) is a metrics based system with peer oversight of the design of a package of indicators and monitoring of their use. RAE2008 could be used to help develop a set of appropriate discipline based metrics, looking at all parts of the research process, not just inputs. Groups would oversee the selection and validity of the metrics, monitoring and, where needed, interpreting and moderating their outcomes. This could include a mechanism to recognise early stage researchers, ie, making sure that metrics do not focus solely on inputs which would favour those already established, and also potentially through the submission of evidence of how the institution is incentivising and supporting emerging excellence.

5.6 In conclusion, Universities UK believes that time will be needed to consider the full implications of each of the options and that we should work closely with the funding councils to develop a new mechanism. This will ensure that a cost effective metrics-based system balanced by appropriate peer involvement is agreed as a sustainable successor to the Research Assessment Exercise.

Annex A:

The future of research funding and assessment: Universities UK's key principles

Dual support

- Any new system should be consistent with the Government's commitment to the continuation of the dual support system.

Universities UK agrees with the Government that dual support is an effective mechanism to sustain research excellence. A mechanism that directly relates all funding council research support to the income earned through grants and contracts can no longer be regarded as dual support.

Quality Research (QR)

- Any new mechanism driving QR allocations should be appropriate to the uses that QR is put within a dual system.

Any new system should support the continuation of the unhypothecated nature of QR, which would be consistent with the block grant principle and the effective operation of the dual support system.

Quality assessment and peer involvement

- Funding allocations should be selective and based on a judgement of quality, with peer involvement.

The current process consists of two distinct but related components: internationally benchmarked quality assessment, which then informs the allocation of funds. The Government's current proposals only deal with the second of these. Some form of quality measurement needs to continue and should not be separated from the funding allocation method.

Metrics will have a key role to play in any new mechanism. However, most of the metrics suggested to date are proxy indicators of quality, rather than measures of quality. Therefore, expert assessment (on a light touch basis) must continue within the process, for example, by overseeing the selection of the metrics and monitoring and interpreting their outcomes.

Universality

- Any new system should continue to be UK-wide and provide universal coverage across discipline areas.

It will be important to have a review mechanism that provides robust quality measures, which each of the devolved funding councils can subsequently use to allocate funds on the basis of their own policies. This will also ensure that a single, UK-wide, system is available for international benchmarking and to promote the capability and reputation of the UK research base. Equally, all subject areas and institutions must be incorporated within the same assessment framework (though there may be differences of detail in subject areas as there are in the current system).

Applied and user focused research

- An appropriate and effective mechanism is needed within any new system to ensure that applied and user focused research is recognised and supported.

There is a need for a more explicit recognition of this type of research. Metrics in this area will need to be robust indicators of quality as well as quantity.

Behavioural impact

- Any new assessment system should seek to limit the deleterious impact on institutional behaviour.

Systems that measure and reward performance will affect behaviour and they are intended to do so. It will, however, be important when designing a new system to be aware of and minimise negative distortions.

Specifically, any new system should:

- seek to limit the incentives for institutions to simply 'chase money';
 - seek to limit the incentives which have the potential to distort human resource and recruitment practices.
- Specifically, it should not hamper the development of the next generation of researchers.

Transition and financial stability

- The transition to a new system should be managed and moderated to avoid any destabilising effects. Any new system should also provide a sufficiently stable financial framework that allows institutions to invest and plan on the basis of some reasonable assumptions about future levels of income.

Universities UK believes that the outcomes of RAE2008 should be the primary quality mechanism informing funding for at least a four year period. This will ensure that there is relative stability in the system, and that confidence in the coming exercise is maintained. It will also provide sufficient time to develop a robust alternative and manage the transition between the two systems.

Cost and simplicity

- **Any new system should seek to reduce the administrative burden on HEIs and the effort involved in transforming the system will need to be proportionate to the gains.**

The Government's intention to reduce the burden of research assessment is welcome and it will be vital to undertake a thorough assessment of the likely impact of any new model.

Policy fit

- **There needs to be a more thorough understanding of how the Government proposals for reform align with, and impact upon, other policy initiatives and objectives.**

Universities UK has some concern that debates around reform of research assessment and funding are taking place within a 'policy vacuum'. Other areas where the implications need to be understood include implementation of full economic costing and the Higher Education Innovation Fund (and equivalent knowledge transfer funding streams in the devolved administrations).

Annex B:

Research assessment: main options for reform

Option A:

Metrics-based assessment

Income based

Determined by
Government

Option C:

Metrics-based assessment

Peer review trigger (based
on significant variation in
performance compared
with the RAE2008
baseline)

Option B:

Metrics-based assessment

Multiple indicators

Peer determined and
monitored

Option D:

Peer review

Streamlined version of the
current exercise supported
by panel use of metrics

Notes:

1 *Department of Trade and Industry: departmental report 2006*, HMSO, London.

2 Geuna A & Martin B R *University research evaluation and funding: an international comparison*, *Minerva*, Vol.41 (2003) pp.277-304

3 HEFCE evidence quoted in 2003-04 House of Commons Science and Technology Committee report *Research Assessment Exercise: a re-assessment*

4 *Ibid*

5 Royal Society (2003) *Supporting basic research in science and engineering: a call for a radical review of university research and funding in the UK* & House of Commons Science and Technology Committee *Research Assessment Exercise: a re-assessment: eleventh report of session 2003-04*, TSO, Norwich

6 Geuna A & Martin B R, *op cit*

7 Undertaken by Professor Andrew Walker, Heriot-Watt University

About Universities UK

Universities UK is the representative body for the executive heads of UK universities and is recognised as the umbrella group for the university sector. It works to advance the interests of universities and to spread good practice throughout the higher education sector.



Universities UK

Universities UK

Woburn House, 20 Tavistock Square, London, WC1H 9HQ

Tel +44 (0)20 7419 4111 Fax +44 (0)20 7388 8649

Email info@UniversitiesUK.ac.uk Web www.UniversitiesUK.ac.uk