Joint funding bodies' review of research assessment
De Montfort University Response

De Montfort University is the leading post-1992 university for research on the basis of the 2001 RAE. The University conducts research across a wide range of disciplines with different cultures and expectations. This response is based on full consultation with staff representing these internal constituencies. The University fully supports the review:

Group 1: Expert Review.

We endorse the principle of peer review. It has been one of the key features of previous research assessment exercises and has inspired confidence in the academic community. However, in areas where 'peers' are not necessarily synonymous with 'experts', e.g. science and engineering, we recommend the increasing use of experts (e.g. from Research Institutes, Industry or from overseas) in order to provide objectivity and balance. We do not believe that research and teaching assessments should be combined.

a. Should the assessments be prospective, retrospective or a combination of the two?

Assessments should be a combination of the two. A track record of achievement is in many instances the most reliable indicator of continued future development. This track record should not be confused with performance in previous assessment cycles, which is no guarantee of future achievement. Retrospective assessment should be complemented by a statement of future plans to indicate the dynamism and the upward mobility of different departments. Statements should contain a strategic thrust plus evidence of how plans will be supported.

b. What objective data should assessors consider?

Data should be output-led. We recognise that these outputs can vary by discipline and that data other than publications might be less significant in the Humanities than in Science and Engineering. Data to be considered should include:

- Publications
- Research grants from research councils
- Research grants and contracts from other sources including those that relate to the applications of research and its commercial purposes
- Numbers of doctoral students and completions
- Numbers of research-active staff in the institution, but not as proportionate to the staffing base as a whole
- Indicators of peer esteem, including citations, prizes, awards, membership of learned societies, conference invitations etc.

We would welcome a system that provides clarity as to how elements such as 'research culture', esteem indicators etc are weighted in relation to the outputs and in relation to metrics such as research income and doctoral students.

Similarly we would welcome a system that provides clarity regarding the position of young or new researchers. We would argue strongly that a key element in a good research ethos is the development of young researchers and that a more differentiated approach to staff submission could be adopted. Such a differentiated approach is consonant with a research development culture based on team work, project research and co-authoring, and it would be important for any new system to recognise and incentivise such an approach.

c. At what level should assessments be made – individuals, groups, departments, research institutes, or higher education institutions?

Assessments should be made at the level of research groups formed by combining individual data. Groups would normally, but not automatically, be located within a departmental structure. Assessment must be sensitive to the differential structures for research operating within the HE sector.

Institutional ownership of the submission is also vital. Research funding supports the research infrastructure (selectivity notwithstanding). Institutional ownership of the submission requires the organisation to accept responsibility for the development of its research and for its future efficient exploitation.

d. Is there an alternative to organising the assessment around subjects or thematic areas? If this is unavoidable, roughly how many should there be?

Assessment should remain subject based. The number of categories should reflect the present structure but should take into account the critical mass of researchers in any single subject area so that subject panels are broadly comparable. Where there are particularly small units of subject assessment at the moment, these should be integrated into thematic groups of cognate disciplines so that parity is maintained. Any system needs also to recognise the value of research that is inter- or multi-disciplinary, so as to encourage innovation and exploration across conventional subject boundaries.

e. What are the major strengths and weaknesses of this approach?

Strengths:

- The majority of research activity falls within broad-based subject definitions
- Most researchers perceive their academic identity as deriving from their subject
- Subject communities can define assessment criteria appropriate to their discipline
- Subjects are recognisable and remain the cornerstone of HE structures whilst retaining the flexibility to enable cross-disciplinary research to flourish

Weaknesses:

- Unequal assessment between groups who apply similar criteria with commensurately unequal results
- The system is time consuming and expensive
- Too narrow a definition of what constitutes a subject can lead to an equally narrow application of criteria and can have the effect of failing to recognise diversity, innovation or non-traditional methodologies/outputs
- A false divisiveness and potential discouragement of interdisciplinary research, especially where this might also cut across broader subject categories, such as the arts/sciences divide

Group 2: Algorithm

a. Is it, in principle, acceptable to assess research entirely on the basis of metrics?

We do not advocate the sole use of algorithms in the assessment of research. We do recognize that some metrics, in combination with other data are useful in forming judgements, if applied consistently across disciplines.

b. What metrics are available?

Metrics are a crude measure of quality. Those cited should be modified to include quantifiable data only. In particular we believe that reputation surveys are highly fallible and often based on inaccurate or misunderstood data plus inherited bias. The use of algorithms based on metrics is desirable only in combination with expert assessment. There are also subjects with strong professional affiliations, such as Law, where a standard metrics-based approach is inappropriate.

c. Can the available metrics be combined to provide an accurate picture of the location of research strength?

The available metrics must continue to be used with caution, uniformly and in combination with an assessment of the quality of research output based on peer/expert review.

d. If funding were tied to the available metrics, what effects would this have upon behaviour? Would the metrics themselves continue to be reliable?

The metrics would become less reliable and institutions would become more susceptible to games playing.

f. What are the major strengths and weaknesses of this approach?

Strengths:

- Metrics provide reliable and easily verifiable data
- They establish a genuine sense of continuity between assessment exercises across a lengthy period and allow comparisons of development and excellence to be made
- They enable institutions to monitor their own ongoing performance

Weaknesses:

- If a purely metrics based system is adopted there is no mechanism for future planning; all data will be retrospective.
- Metrics are a crude instrument for assessing intellectual product.
- Metrics cannot measure certain key elements of research quality, such as vision, environment or innovation. A metrics-based approach could de-motivate researchers and deter intellectual advancement
- Metrics are of variable validity across disciplines

Group 3: Self-assessment

A degree of self-assessment is already included in the current system, through RA5&6. It does not require departments to bid explicitly for status or support overly ambitious claims. It is difficult to see how a system relying more heavily on self-assessment could be reliably policed.

a. What data might we require institutions to include in their self-assessments?

Both qualitative and metric

b. Should the assessments be prospective, retrospective or a combination of the two?

Self-assessment, as with the current system and indeed any form of assessment, should combine retrospective and prospective elements.

c. What criteria should institutions be obliged to apply to their own work. Should these be the same in each institution or each subject?

There should be certain institutional criteria relating to institution-wide policies, funding and management of research activity. These should be the same in each institution. There should also be a set of appropriate subject specific criteria, many of which would be applicable to a wide range of subjects. These would need to be particularly robust if they were to secure reliability.

d. How might we credibly validate institutions' own assessment of their own work?

Sampling by experts;

e. Would self-assessment be more or less burdensome than expert review?

Less burdensome on the system; more burdensome on institutions

g. What are the major strengths and weaknesses of this approach

Strengths:

- The process of self-assessment would require institutions to engage in a rigorous and continuous self-monitoring system which could also aid strategic planning and research development
- Institutions could identify their own level of achievement within the context of a distinctive institutional mission
- It can help institutions to review and demonstrate their own development
- It could more easily recognise cross-boundary research
- It is less expensive to administer
- It could reduce bureaucracy

Weaknesses:

- It is more difficult to ensure objectivity and consistency.
- It could lead to over-exaggerated claims that could be difficult and time-consuming to verify
- Partial scrutiny of submissions could threaten confidence in the system
- It is inappropriate as a sole determinant given the substantial amount of funding at stake
- It compromises one of the key principles of the present system, the reliance on peer review

Group 4: Historical ratings

We are strongly opposed to the introduction of a system based on historical ratings. Given that not all institutions started from the same or even comparable base lines, we feel that such a system would effectively discriminate against institutions, especially the post-1992 universities that have a remarkably successful record of recent development and that have demonstrated research excellence and distinctiveness with very limited investment. A historical ratings assessment would encourage stagnation in the system.

a. Is it acceptable to employ a system that effectively acknowledges that the distribution of research strength is likely to change very slowly?

The system must be flexible enough to recognise rapid change. Historical ratings, as well as recognising continuous excellence, must also reward beneficially those departments that are shown to have dramatic rates of improvement during the years of the RAE and are worthy of future investment

b. What measures should be used to establish each institution's baseline ratings?

Continued excellence or progress over 3 assessment cycles.

c. What mechanism might be used to identify failing institutions or institutions outperforming expectations? Could it involve a 'value for money' element?

Again, progress or lack of it over 3 (or more) RAEs. A value for money element should be introduced to identify institutions that have been able to develop rapidly with limited investment and conversely to identify those that have been wasteful of resources.

d. What would be the likely effects upon behaviour?

As well as rewarding excellence, any system must also acknowledge mobility and success.

e. What are the major strengths and weaknesses of this approach?

We can see no strengths in such a system.

Weaknesses:

- It could inhibit development in new or emergent areas.
- It would be inflexible and resistant to change across the sector.
- It could discourage productive cross-institutional partnerships.
- It could lead to complacency in departments/institutions with a historic or reputational track record
- It would disincentivise the sector as a whole
- Historical ratings would reward past (sometimes long past) performance and discriminate against emerging areas of research excellence.

Group 5: Crosscutting themes

a. What should/could an assessment of the research base be used for?

Assessment of the research base should be used to ensure sustainable funding for high quality research. It can provide information about the location of research quality and about research trends on a cross sectoral basis. It can assist institutions in planning and development.

It is, however, essential that funding bodies and councils do not link their funding schemes too closely to RAE results as this can block funding opportunities and consequently remove one of the most successful vehicles for mobility. A singular system can invest too much power in too few places and importantly will confuse assessment with support.

b. How often should research be assessed? Should it be on a rolling basis? How often should research assessment take place? Should all subjects and all institutions be assessed at the same time or with the same frequency? Should clusters of subjects be assessed separately?

A cycle of six years would prevent some of the current abuses of the system. It would encourage intellectual vitality and risk-taking whilst retaining a degree of continuity. It would facilitate a publishing strategy by allowing for longer development time for individuals and teams, and could generate a greater diversity of published outputs (e.g. books as well as articles). A cycle of this length would fit better with typical 'input-output' cycles in research, and would allow for the increasing long lead times for publication. A longer time period has important equity advantages, with the potential to reduce the handicap experienced by part-time staff or individuals taking time out during an RAE period (e.g. for maternity, family or sickness leave, but also time on secondment in 'practice' or undertaking major pieces of policy-related consultancy). A longer cycle than at present would reduce the administrative costs.

To lengthen the period between assessments beyond six years could prolong outdated distributions of resources.

All subjects and all institutions should be assessed at the same time. A rolling system could increase the administrative burden.

c. What is excellence in research?

Excellence and quality are elusive terms that evade generic description. Although there are quantitative measures that can be employed to evaluate the amount and the scope of research activity, these cannot evaluate excellence nor do they have the capacity for fine discrimination. Excellence can be defined as the best work in any field as determined by its peer group and its users. Excellent research genuinely advances thinking and understanding and/or has a wide-reaching application with the capacity to change the world in which we live.

d. Should research assessment determine the proportion of the available funding directed towards each subject?

A major problem in using RAE results to determine funds available for different subjects is the evidence that assessment has not been consistent across UoA's. Whilst panels might have been consistent in their own interpretation of criteria, there was significant variability in judgements across cognate subject areas. For example, differing degrees of upwards 'grade drift' were apparent in different UoA's between 1996 and 2001. Although grade inflation is denied by panels in their reports, it is striking that e.g. 62% of submissions to Law (UoA 36) were graded 5 or 5* in 2001, compared with only 16% of submissions in Business & Management (UoA 43). If standards of assessment are not consistent across units of assessment, and thus if RAE is used to distribute funds, some units will be unfairly disadvantaged.

The relationship between assessment and funding consequently needs to be clearly articulated at the start of any assessment exercise.

e. Should each institution be assessed in the same way?

Yes. This is the only sure measure of equity. Processes should be sufficiently flexible to take account of institutional ethos, patterns of development and historical factors. We would not endorse a two tier system of assessment or support.

We support the notion of a ladder of improvement but would require reassurance as to how this would be benchmarked. We would support having targets of improvement so that progress could be embedded in the system.

Account must also be taken of the fact that some universities are much less likely than others to be 'broad spectrum' research institutions and should not be disadvantaged as a result. An assessment system should not discourage institutions from engaging with other government priorities, such as widening participation. Nor should it discourage individuals from writing important text books. We believe that universities can engage in the full range of activities and should not be deterred from their mission.

f. Should each subject or group of cognate subjects be assessed in the same way?

Subjects or groups of cognate subjects should be given flexibility to define their own criteria within indicative guidelines. The process of assessment, however, should be comparable across disciplines to ensure equity.

h. How much discretion should institutions have in putting together their submissions?

Institutions should have discretion in the submission process in order to reflect their distinctive approaches and research ethos. They should also have discretion over which staff they submit.

i. How can a research assessment process be designed to support equality of treatment for all groups of staff in Higher Education?

The current system implicitly discriminates against women or those who have child care responsibilities. A longer assessment cycle would benefit those taking career breaks e.g. for maternity/paternity and childcare purposes.

j. Priorities: what are the most important features of an assessment process?

To be fair, to get it right and to be transparent. 'Fairness to individuals and institutions' – in practice and in perceptions – should be the dominant criterion for an assessment process. The process should also aim to minimise the costs and bureaucratic burdens so that the funding can be properly directed towards the support of the research activity itself.

Group 6: Have we missed anything?

All research councils cite the desire and needs for a benefit to UK plc and yet there appears too little effort to quantify this contribution. It is recognised that such an assessment would be difficult but one approach might be via patent registration followed by a measure of the industrial licence agreements based on those patents.

Reference to the interdisciplinary debate is a surprising omission, especially in the context of recent reviews of and developments in research across the traditional arts/science boundaries.

There is also no reference to the wider context of HE. Research cannot be isolated from the current funding crisis and the related governmental confusion about the purpose of Higher Education. This context needs to inform the review of research assessment across the sector.