UNIVERSITY *of* ABERTAY DUNDEE

Response to Joint Funding Bodies' Review of Research Assessment

1. The University welcomes the opportunity to respond to the Joint Funding Bodies' initial consultation on the review of research assessment. The commentary below represents the views of the University's research community.

Introduction

2. The main driver for the Research Assessment Exercise is the further development and maintenance of excellence in research. The University fully supports this policy and the evidence of past RAEs suggests that the exercise broadly delivers the required objectives. To this extent the University does not see a need for a radical change in the overall mechanisms of, or philosophy behind, the RAE, but does consider that there are issues which the review process should confront.

Approaches to assessment

Expert review

3. There is no doubt that any form of assessment will need to retain some form of expert review if it is to remain credible and maintain consistency between assessments. Should this be in the form of panels, the membership of a panel should be representative of the sector as a whole, and where appropriate (and much more so than in RAE 2001), contain non-academic experts, e.g. end-users from industry, evidence-based practice etc., and lay members, i.e. people who can assess the significance of the work to beneficiaries outside the sometimes rather limited confines of the academic community. An increased use of experts from outside the UK may also be beneficial, given that the quality of the research is to be assessed in an international context. Additionally, should panels remain discipline focused, a significant proportion of the panel membership should be made up from those who work at the interface with other disciplines or even, where appropriate, in other disciplines; this would ensure a more consistent appraisal of inter-disciplinary work.

4. The criteria for panels should be published as far as possible in advance of any assessment deadline and care should be taken to ensure that they do not encourage a risk averse approach to developing research. Each panel should have an independent Chairperson/administrator with the authority to ensure that panels do not drift outside their stated criteria. Panels should provide detailed feedback on the rationale behind their assessments; and in rare cases where this feedback suggested that the panel had not followed their stated criteria in making their decision, an independently controlled appeal process could be operated. Alternatively, panels may wish to provide comments/questions/feedback before making their final decision on the grade; submitters would then be able to respond to the points raised by the panel in a similar way to that in which applicants are often given the opportunity to respond to referees' comments on a research proposal. Given that the experts will be leaders in their field, provision could also be made for them to provide guidance on what steps could be taken to further the development of research quality in the areas assessed.

5. The assessment should assess both past performance *and* potential future performance. The assessment of future potential is vital for the assessment of new and emerging areas with no tradition of assessment. Where, for example, new research groups are established in novel areas, or areas in which an institution has no research tradition (often representing a significant investment in infrastructure), in the period preceding an assessment deadline, they are unlikely to have sufficient appropriate outputs to submit to the upcoming RAE. Assuming they submit to the following RAE, and that RAEs continue at regular intervals, this means that they may have to wait 6, 7, or even 8 years before they qualify to receive funding from the Funding Council. We suggest, therefore, that such groups be able to submit on the basis of their potential, rather than track record, in a particular area; this could take account, for example, of the abilities of the individual staff appointed (track record in other areas, whether in an academic or non-academic (e.g. industry, NHS) environment), the investment in infrastructure by the institution and external organisations, recently awarded project funding, and the plans against which future performance can be evaluated (note

the parallels to research grants, which are also awarded largely on the basis of the quality of *future* research plans). It may be desirable to separate the assessment of such groupings from those able to meet the normal submission requirements (in terms of outputs) for an RAE. The results of the assessments based solely on potential could feed into a separate 'development' funding stream, such as the Research Development Foundation Grant currently operated by SHEFC. In this way, the assessment could serve a funding structure which integrates core funding with a forward-looking allocation to accommodate adventure and initiative in research.

6. The data to be considered should be the usual measures of research quality used in previous RAEs (including future plans/potential), perhaps with additional measures relating to knowledge transfer (patents, licensing, spin-outs etc.)[1] and a measure reflecting the increase or decrease in performance relative to previous assessments. A 'value for money' or 'value added' indicator would be helpful in identifying areas which were out-performing or under-performing in relation to expectations. This might involve some assessment of their research capability in terms of research infrastructure. However, this might best be achieved by using a measure of efficiency, e.g., relating quality of outputs to the level of input. Those awarded very significant amounts of funding should be expected to do well, and continue to be rewarded for doing so; but be penalised for inefficient use of resources. Those receiving small sums who use them well, should be rewarded for doing so. Such a measure, would, of course, be dependent on the measures of output reflecting the diversity of time-scales (long- vs short-term studies) and outcomes of research activity. The value for money or value added indicator could either be incorporated in the final grading, or given as an additional weighting, which Funding Councils may wish to use when distributing funds.

7. A 'value added' indicator might also take account of a number of other measures. The importance of research to teaching could be emphasised if researchers who teach their subject to substantial numbers of undergraduates gained some credit for knowledge transfer; for example, researchers could be given credit for teaching in the areas they publish in. Fifty per cent of those completing PhD's leave science within two appointments; practices designed to develop and retain the best researchers, particularly contract researchers, in research, could be rewarded. Employment profiles could be used to determine whether PhD students and contract researchers leave for positions that utilise their research training/expertise; good supervisors could be given credit. Researchers that share their knowledge with students, post-docs, industry, public bodies and the community could be given credit. Research is certainly not just about the production of papers.

8. With the increasingly inter-disciplinary nature of research, with much research taking place at the interface between 'disciplines', assessment at 'Department' level has little meaning in many areas. In many cases Departments are simply convenient administrative units, with research 'groups' cutting across many Departments, Schools or Faculties, often addressing a variety of themes. In this respect it is rather meaningless to denote a 'level' at which assessments should be made. Universities should be free to submit appropriate individual scholars (particularly where no related research expertise exists in the institution) or groupings (which may equate, e.g. to 'Departments' in some cases, or inter- or multi- disciplinary groupings in others) which they believe have demonstrated research excellence in a particular area, or in the case of those initiating research in new and emerging areas, have the demonstrable potential to do so (see paragraph 5). However, assessment must not be at the level of the HEI, since a research assessment exercise must have as its primary goal the *recognition* of excellence. Assessing the rapidly evolving new universities (which have not traditionally been resourced for research) on this basis will mask 'pockets of research excellence' and niche areas in which newer institutions are often at the forefront, thus disenfranchising a large segment of the research community. Ultimately, this will inhibit the emergence of new areas and disciplines which have no research tradition, and the UK's ability to compete globally in a diverse research environment.

9. It is hard to see how the assessment could operate without grouping research around subjects or thematic areas. The key change should be to ensure that a proportion of the experts making up any panel or other reviewing system are able to judge the worth of inter-disciplinary submissions within these areas. Institutions could also be asked to 'flag' inter- or multi- disciplinary submissions which they believed had cross-panel relevance. An additional possibility would be a new Unit of Assessment for the evaluation of any research

[1] We are aware that a study of the most appropriate metrics for Knowledge Transfer is already underway in Scotland.

which did not fit readily into one of the current discipline specific Units. Perhaps there should even be a generic 'Industrial Research' Unit of Assessment?

10. The strength of expert review lies in its general acceptance by the academic community as the most appropriate method so far devised for the assessment of research (grant applications, journal articles etc.), although it inherently favours established areas and groups. Criticisms most commonly arise when the review process is perceived to be 'clouded in mystery', inconsistent with stated criteria, unrepresentative or biased towards a particular section of the community ('who you know' rather than 'what you know') or overly conservative; in the latter case this may stifle the development of more adventurous or risky research in favour of the safe and familiar. It is therefore vital that any expert review process used for research assessment is transparent and fair, with clear criteria and experts who can be seen to be representative of the academic community as a whole, and of appropriate non-academic end-users.

Algorithm

11. Basing research assessment *solely* on an algorithm would be unacceptable. Unless a 'research potential' metric could be introduced it would effectively restrict recurrent funding to established areas, with no way forward for new and emerging areas. It would undoubtedly reduce the diversity of the research base since researchers would be tempted to work only in areas which produce 'quick' results, and hence publications, and to work only in the most popular fields or jump on the latest 'bandwagon' to ensure citations, rather than move into new areas or work in areas which by their very nature require a longer-term commitment before meaningful results are obtained. We should not discourage those who wish to stimulate the development of new areas, work on longer-term problems or undertake more adventurous research, by an over-reliance on metrics such as bibliometric measures. The purpose of citations is to refer readers back to essential references; if the number of citations received ultimately has a significant influence on the funding awarded, there is a real danger of encouraging a culture of sterile mutual citation.

12. Using research student numbers or completions (unless as a ratio of starts/within 4-year finishes) as a major factor in an algorithm would effectively block the development of Departments rated below 3a in England, should HEFCE's proposals to remove funding for postgraduate students in such Departments go ahead. Measures of financial sustainability are problematical, given that sustainability may be heavily dependent on the outcome of the assessment exercise itself, and the consequent impact on recurrent research funding.

13. The major weakness of this approach is that it would effectively be based on past funding history. It is to be expected that research areas and institutions that have historically received the most recurrent research funding will have evolved to a stage where their infrastructure, staffing and research quality, and hence their ability to attract the most external income and research students, is reflected in a high level of outputs of the type that would make up the metrics used in an algorithm. (e.g. publications, licences, spin-outs). Indeed, if this were not the case, we should be asking why? Other, newer, research areas and institutions, having received comparatively little recurrent research funding, are at a relatively early stage of their evolution in research terms and hence produce a lower level of outputs of the type that would make up the metrics used in an algorithm. However, such research areas and institutions often 'punch above their weight' with little reward. What is needed is a measure of efficiency, e.g., relating the quality of outputs to the level of input (see paragraph 6).

14. The strengths of the algorithm approach are that it would be a relatively simplified, transparent and less expensive and time consuming process (providing the required metrics were defined well in advance of the first assessment). Principal weaknesses of this approach are that it takes no account of research potential and people would work to the metrics, thus stifling innovation.

Self assessment

15. Self-assessment is an attractive proposition since the institution could timetable reviews to fit with its strategic development plan. Areas to be reviewed could be selected on a rolling basis (instead of a one-off big-bang), allowing the institution to place a developmental focus on its research activities. Research judged to have attained a particular level of excellence by internal review could then be submitted for external expert review for validation/audit (cf. QAA). Such a situation would benefit rapidly evolving areas of research and institutions since their improvements would be recognized and incorporated into funding allocations

without the time-lag experienced with the current one-off assessment exercise. Areas of research assessed internally to be relatively stable could continue to be submitted for external validation at set intervals. The assessment should assess both past performance *and* potential future performance for established areas, and future potential for the assessment of new and emerging areas with no tradition of assessment.

16. It is likely that self-assessment would be less burdensome than expert review, since many universities already conduct some form of self assessment. Moreover, in conjunction with external validation and audit, the process will provide universities with reliable management information and facilitate strategies for the future development of the institution's research portfolio.

Historical ratings

17. A system based solely on historical ratings would be wholly unacceptable. It would result in a static system entrenched in the past, rather than a dynamic system able to respond to the needs of a rapidly changing world. It would cement the divide between long-established institutions who have built their reputations and track record with the benefit of high levels of historical funding and the rapidly evolving 'late starters' who have made significant progress in a short time. Such a system would reduce the diversity of the research base, since the newer institutions are often most active in niche areas and in areas relevant to non-academic end-users such as industry. Whilst it may enable those institutions in receipt of the majority of funding to plan future developments on a more secure basis, it may also lead to complacency in research and inefficiency in the use of funds.

18. A 'value for money' rating is important if we wish to ensure that research resources are used in an efficient manner. As previously mentioned (paragraph 6), this might best be achieved by using a measure of efficiency, e.g., relating quality of outputs to the level of input (recurrent and external income, research infrastructure already in place). Such an indicator is likely to be of most value in evaluating resource intensive research areas. However, it is not the case that a 'value for money' indicator could only operate within a system based on historical ratings; it could be incorporated into almost any system of assessment, or used to adjust the funding awarded (as a major, not minor, indicator) at the funding allocation stage.

19. A value for money indicator may encourage institutions to take a more focused approach to distributing the RAE-based grant; whilst this is likely to have a positive effect in many cases, if it were based simply on targeting areas likely to produce the most outputs, citations etc. it may discourage investment in more adventurous or time-consuming research, thus reducing the diversity of the research base. A better balance might be achieved if some form of 'value added' indicator were also used (see paragraph 7).

Cross-cutting themes

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

20. For Funding Councils the primary function of research assessments will continue to be to provide a mechanism for distributing limited funding. From a researcher's and institution's point of view the purpose of the assessment should be to provide a fair and transparent method of recognizing and rewarding research excellence and potential *wherever* it is to be found. In this context, all institutions should be assessed in the same way. From an institution's perspective the more reliable management information and quality feedback provided by the exercise the better; this will enable more informed strategic management and development of an institution's research portfolio. It is certainly important that the Funding Councils should explain what the information produced by the exercise means, and what it should be used for; this might discourage the use of such information out of context, as in the construction of league tables.

21. Serious consideration should also be given to specifying what the outcomes of a research assessment exercise should not be used for. In particular, they should not be used as proxies for other funding decisions. For example, they should not be allowed to influence in any way the outcomes of research grant applications or (unless technology transfer metrics are included in the assessment) funding to support commercialisation.

How often should research be assessed?

22. Some form of rolling or continuous assessment process is attractive since it would allow institutions to submit new and emerging areas for assessment, or rapidly improving existing areas for re-assessment, without the long time-lag experienced under the current system. This would encourage institutions to place a developmental focus on their research activities. More stable areas could be assessed at longer time intervals. If a rolling assessment process is not used, then 5 year intervals are reasonable for one-off assessments.

What is excellence in research?

23. Defining 'excellence in research' is not easy, as is widely accepted. In practice, excellence in research is what one's peers deem to be excellent. However, such peer review tends to rate 'scholarly' activity or basic research higher than applied or practice based research. Additionally, research that is ahead of its time, or doesn't agree with current theories, can be dismissed. This can discourage adventurous research. A more appropriate description of excellence in research (assuming appropriate and rigorous methodology) might be that which has the greatest impact; either on the theoretical development of a discipline or topic, on the provision of solutions to practical problems and the key issues confronting human society, or in the creation of something genuinely new. Certainly, creativity, applicability and value to the end-user (whether academic or non-academic) are aspects that deserve more recognition than was the case in RAE 2001.

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

24. The distribution of funds between subject pots would most appropriately be decided post-assessment. The decision could incorporate all of the measures stated, although historical distribution should be seen as the least important.

Should each institution be assessed in the same way?

25. The exercise aims to assess quality. Clearly then, the quality of the research should be assessed in the same way, independent of the institution. For example, one would expect consistency within the assessment process such that an output assessed to be of high quality which was submitted by a researcher from a large institution with a strong research tradition, would receive the same quality rating were it submitted by a researcher from a smaller institution with no tradition of large-scale investment in research. Similarly, with self-assessment, although there would need to be some flexibility in approach, to reflect e.g., different organisational structures, the criteria regarding research quality would remain the same, with all institutions subject to external validation to ensure consistency across the sector.

26. However, within the assessment process (or even post-assessment) there should be scope for recognizing the developmental nature and potential of research in newer universities, with the aim of encouraging development in areas of potential rather than punishing institutions for being 'late starters'. Researchers in newer universities are often operating under greater constraints than those in their longer established counterparts, with a comparative lack of research infrastructure, greater teaching loads, and fewer research active colleagues. They may be working as individual scholars or in small groups, in larger inter-disciplinary groups (portions of which are submitted to different Units of Assessment and hence hide 'critical mass'), or in collaborations with researchers from other institutions. It is important that, where such researchers are producing high quality work, they are not penalised because they are deemed to be working in a lower quality environment (infrastructure, facilities) with a perceived lack of 'critical mass'. On this point there is currently such emphasis on 'critical mass' that one could be forgiven for thinking that lone scholars or smaller research groupings are no longer capable of producing high quality or internationally relevant research; this is clearly not the case. It can be argued that the production of high quality research by researchers operating in institutions or research areas that have received little investment is more meritorious than the production of similar quality research in institutions or comparable research areas that have received substantial investment over the years.

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

27. Some form of overarching framework should be provided for each subject or group of subjects to maintain consistency between related areas. However, disciplines/subjects vary widely, and hence a 'one size fits all'

approach is not suitable; it may well be appropriate to allow broad groups (e.g. Arts and Humanities, Sciences) to diverge significantly in their assessment criteria. Confidence in the system would be greatly increased if the academic community were first consulted on the most appropriate criteria/outputs to be used in their disciplines/subject areas; four papers is somewhat arbitrary and may not be appropriate to some disciplines. Even within disciplines, it should be recognized that the answers to some research questions can be obtained relatively quickly (leading to rapid publication), whereas obtaining the answers to others requires a much longer term commitment (with associated knock-on effects on publication rates).

28. It is also important to ensure that inter- and multi- disciplinary work can be assessed appropriately across areas which may be using different criteria; perhaps a number of specific inter-disciplinary panels or sub-panels are required (whether as per RAE 2001 or for validation of self-assessment)?

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

29. Institutions know where their research expertise lies and should be free to demonstrate it; they should therefore have the maximum possible discretion over which staff should be assessed. The two alternative systems suggested in the consultation document are not mutually exclusive; institutional decisions would be expected to take account of the views of individuals, research groups etc.

30. The assessment exercise should be designed to promote and recognize research excellence, not to ensure that all staff in all universities are engaged in research. However, if it is decided that all staff should be submitted (to eliminate 'game playing' or perceived unfairness to research active individuals whose work might not otherwise be included), then institutions must be allowed to discriminate those that are research active from those that are 'research inactive'; the latter (e.g. staff employed for teaching only) could be subject to random audit to ensure that they were indeed inactive in research. We should be assessing the *quality* of the research that *is* done (or will potentially be done, for new groups/areas), not becoming preoccupied with what is not done (although we should not underestimate the contributions of those who free up time for others to do research by taking on more teaching and administrative duties). On no account should the assessed excellence of the active researchers be diluted by calculating an average using the total number of staff (active + inactive) employed at the institution. Such a system would serve only to hide research excellence, not to recognize it. This is particularly important for research active staff employed in the modern universities, where a significant proportion of the established staff were initially employed only for teaching.

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

31. Clearly, certain categories of staff will be expected to be included in the assessment process. However, a pertinent design feature of a research assessment process that encourages genuine equality of opportunity for all is that *any* staff member conducting research in an institution should be eligible to be included in the assessment process, regardless of their status/title, if they are research active as an individual. This is particularly relevant in relation to post-doctoral contract researchers, who do a significant amount of the research assessed, but (under the current assessment system) with little recognition of their contribution. How universities respond to the new employment legislation for contract workers will clearly influence this area.

32. Given that the next assessment is not expected to take place until 2008/9, provision should be made to ensure that the contributions of staff who retire before the exercise can be included. It is also important that the system enables proper account to be taken of 'time out', where researchers have taken a career break, for example through childcare/family commitments or illness (e.g. these staff should not be penalised for submitting proportionately fewer outputs).

33. The research assessment process will be discriminatory against individuals or groups of staff if their research excellence is not recognized simply because they are working in an institution that does not have a strong tradition of research. The danger is that their excellence will be masked if their contribution is averaged with that of research inactive or newly appointed staff.

34. The RAE as it stands discourages the employment and development of younger researchers. The question on many employers' lips when appointing new staff is 'are they returnable to the RAE?'. It should be 'what is their potential?' Any new assessment process should aim to nurture the next generation of researchers, not discourage and discriminate against them. The current over-reliance on track record for peer-reviewed grant funding is demoralising enough for young researchers, without the RAE making it worse. Submitting new, young staff on the basis of their potential should therefore be seen as a positive in any new research assessment exercise, not as a negative (risking a lower grade) on the basis of insufficient publications in the assessment period or limited track record.

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

35. The most important characteristics of an assessment process from the list provided are transparency and fairness to individuals and institutions. The assessment process should recognize that research is done in institutions *by* people; in this respect it needs to ensure that the system recognizes excellence *wherever* it occurs and provides a framework for encouraging and nurturing the next generation of researchers.

36. Particular care should also be taken to ensure that any new research assessment exercise does not encourage conservative approaches to research; this may arise where researchers are unwilling to pursue any line of research (e.g. inter-disciplinary) which is not thought to be sufficiently aligned to the stated assessment criteria, and hence which they perceive may prejudice their grade. The criteria should be flexible enough to encourage research diversity, rather than constrain it.

37. Clearly, any research assessment process must also be cost effective. Costs should be considered in the broadest sense, not just in terms of the administrative costs to HEFCE of running the process. There is clearly also a large bureaucratic overhead on institutions, with the demands placed on smaller institutions being proportionately greater, especially when considered within the context of the likely gains. However, there are also hidden costs. The institution may be missing out on other opportunities by devoting such a large amount of effort to the assessment process. There is also the cost (loss of output) of taking leading researchers out of the research production loop and, as expert reviewers, into a backward-looking audit loop. The sophistication and cost of any assessment exercise therefore needs to be balanced against the likely gains for all participants (Government, Funding Councils, institutions, individuals).