Analysis of responses to the Research Excellence Framework consultation

Background

1. In November 2007 HEFCE issued a consultation paper on the assessment and funding of higher education research post 2008, inviting responses by 14 February 2008.[1]

2. We set out proposals for the Research Excellence Framework (REF), with the following broad features:

- The REF will be based as far as possible on quantitative measures. There will be an overarching framework within which differences between the disciplines will be accommodated. For the science-based disciplines, funding and assessment will be driven by bibliometric indicators of research quality and data about external research income and research students. For the arts, humanities and social sciences, there will be a light touch peer review process, informed by metrics.

- Funding and assessment will operate at the level of six or seven broad subject groups covering the sciences, engineering, technology and medicine; and a larger number of subject groups for the arts, humanities, social sciences and mathematics and statistics. The process will be overseen by a panel of experts for each subject group.

- For the sciences, the new framework will be phased in gradually from 2010 until all our research funding is driven by it from 2014. For the other subjects, the light touch peer review process will take place in 2013, to drive funding from 2014.

3. For further details of the proposals, refer to the consultation document (HEFCE 2007/34). The proposals were informed by two consultancy studies on the construction and use of bibliometric indicators of quality; further details of these are available on the HEFCE web-site under Publications.

4. We received 274 responses to the HEFCE consultation paper (338 including responses sent to the other funding bodies) including from all of the relevant major national representative and stakeholder bodies, most HEIs and many subject bodies. A list of respondents is at Annex A.

[1] 'Research Excellence Framework: Consultation on the assessment and funding of higher education research post-2008' (HEFCE 2007/34). Parallel consultations were initiated in Wales and Scotland, on the basis that the UK funding bodies would wish to develop a common UK wide system for research quality assessment, which each funding body could use in funding allocations as they saw fit.

5. We also held a series of consultation events for HEIs, and discussed the proposals at a number of other meetings and events with a range of stakeholder groups.

6. This document provides an analysis of the consultation responses from across the UK, and takes into account key points raised at the events. It sets out the key points, followed by a summary of responses to each set of questions in the consultation document; some of the issues overlap and are raised in more than one section.

Key points

7. A number of key stakeholders prefaced their responses by expressing strong support for the dual support system of research funding, stating that the quality-related (QR) element of funding provided as a block grant by the funding bodies is vital for the health and success of the UK research base.

8. A number of responses welcomed or agreed with the proposed aims of the REF, in particular:

- There was strong agreement that the REF should promote excellent research wherever it is found; that it should allocate QR selectively on basis of quality; and that it must therefore be based on robust measures of quality.

- Many welcomed the intention to reduce burden, although often felt there is a tension between reducing burden and maintaining the rigour and robustness of the assessment process.

- There was unqualified support for a UK wide system of quality assessment.

9. Although there was support for greater use of quantitative indicators in the assessment and funding of research, there were widespread concerns about the likely implications of drawing so clear a distinction between two main groups of subjects – science-based, and arts, humanities and social science disciplines – with a single and distinct approach to quality assessment for all subjects falling into each of these two groups. In particular:

- Respondents pointed to the implications of making the division so sharply, especially in cases where it was not clear to which group a discipline should belong or where a body of work was likely to straddle the boundary; and to the implications of applying the same assessment approach without variation to the whole of the diverse range of disciplines within each of the two groups.

- Many respondents expressed a wish to see a more unified system that varies the use of quantitative indicators and expert review, as applicable to each discipline. Some described this as a spectrum or continuum, ranging from ‘metrics moderated by experts’ on one side, to ‘expert review informed by metrics’ on the other, with a graduated range of more mixed approaches in between.

- Concerns were expressed about the operational difficulties of running two distinct assessment systems, including the incentive that this might offer for unhelpful game-playing at the boundary.

10. A clear majority of respondents expressed doubts as to whether a robust and workable assessment regime for science-based disciplines can be developed, piloted and implemented within the timetable that we have proposed. Many respondents requested that we extend this by at least 12 months, stressing the importance of conducting a thorough pilot with sufficient time to refine the proposals before implementation as well as the need to give HEIs sufficient notice of the operational detail of the new assessment regime before putting it into practice. Many also wished to see the timetables for the science-based and other disciplines more closely aligned, to strengthen the development of the new system as a cohesive whole.

11. While views about the proposed bibliometric indicator ranged from positive support to scepticism, there was consensus that:

- Substantive further work is needed to develop the bibliometric process, validate the outcomes and understand the behavioural consequences.

- Citation measures can provide a useful proxy indicator of quality, but they should be considered alongside other quality-related indicators rather than used alone.

- Bibliometric indicators are unlikely to be free from distortion, and it was felt that some form of moderation by expert panels would be necessary to provide confidence in the outcomes.

12. Respondents made proposals for other quantitative indicators that could be used alongside bibliometrics. Most were keen to find ways of capturing user value and the impact of research, although a significant minority felt this is best addressed by other funding streams, and there was little consensus about how in practice impact could be measured.

13. Concerns were raised about the implications of operating at broad subject level for the science-based disciplines. Many felt this would limit the REF's usefulness in informing institutional research management, and in providing public information to external stakeholders about where and in what disciplines excellent research takes place. Many also felt it would also constrain the expertise within the panels. However, a significant minority argued that the REF should focus solely on allocating QR, and not seek to provide detailed quality assessments in the same way as the RAE. They argued that HEIs are now well placed to take more formal responsibility internally for research management without relying on external assessments, and that a more broad-brush approach is necessary for the burden to be reduced significantly.

14. There was consensus that the REF is unlikely to reduce burden in the short term and is expected to involve a significant transitional workload, but many recognised the potential for reduced burden once in a steady state. Many – but not all – respondents urged us to consider a universal or automated approach to identifying staff to be included in the bibliometrics exercise (rather than selection of staff by institutions) and saw this as a major opportunity to reduce the burden on HEIs. For the other subjects, there were general doubts that the burden of peer review could be significantly reduced without compromising rigour, although some avenues for further exploration were identified.

Summary of responses

Subject issues

15. We invited responses to the following questions:

> Consultation question 1a: Do you endorse our proposals for defining the broad group of science-based disciplines, and for dividing this into six main subject groups, in the context of our new approach to assessment and funding?
>
> Consultation question 1b: Are there issues in relation to specific disciplines within this framework that we should consider?

16. Most respondents discussed the implications of the split between science-based and other disciplines. Many had concerns, including the following:

- There was widespread desire for a more unified system that varies the use of quantitative indicators and expert review, as applicable to each discipline. Some described this as a spectrum or continuum, ranging from ‘metrics moderated by experts’ on one side, to ‘expert review informed by metrics’ on the other, with a more mixed approach in between.

- A number of respondents argued that there are substantial differences between the disciplines on each side of the divide (for example, there are substantive differences between engineering and natural sciences, and between mathematics and statistics and the arts and humanities). They argued that a standard approach on either side would not be workable; and noted that the differences within each group are greater than the differences between some of the disciplines on opposite sides of the divide (for example, computer science and mathematics).

- There was general concern about inhibiting interdisciplinary and new and emerging research areas across the boundary between the sciences and other subjects, since such research was seen as increasingly important for addressing major societal challenges (such as aging populations, energy and climate change). There were similar concerns about work between the sciences and other subjects (such as work at the interface between medical and social sciences, statistics and the sciences; biological and medical science related to policy development).

- A number of intermediate disciplines were identified (such as Geography, Mathematics, Archaeology, Architecture, Sports science, Nursing and Allied Health Professions) that do not fit neatly on either side of the boundary. We had proposed that individuals should be assigned to either side of the boundary, and this raised concerns that these intermediate disciplines would fragment or become divided within institutions and departments; and that departments in these areas would have to manage two distinct processes, with differing implications for staff, which would be unduly complex and burdensome. As an alternative, some respondents argued that each discipline should be assessed as a whole under one system or the other, rather than assigning staff on an individual basis. This then raised questions about whether the default for mixed disciplines should be light touch peer review – leaving relatively few disciplines on the metrics side – or whether institutions could choose for these disciplines to be submitted on either side – leading to lack of consistency and coherence across the sector.

- There was concern about the scope for game-playing at the boundary, especially when the initial choices about who goes into the science side would have to be made in advance of developing the criteria for the light touch peer review.

- In general it was seen as inefficient to operate two distinct systems.

- Several respondents commented that describing the two approaches in terms of science-based and other disciplines is unhelpful, and urged that they should be described in terms of different modes of assessment.

- Some felt that a sharp division would reinforce the perception of research in the sciences as being more important than in the arts, humanities and social sciences.

17. Many respondents provided general comments on the proposed subject structure for the science-based disciplines, and discussed the implications of having six broad subject groupings. A number of respondents recognised that broad groupings are suitable for the purpose of allocating funding, and some agreed that they are also suitable for metrics-based assessment. However, many also queried or raised concerns about this:

- Some argued that broad groups are too diverse to be catered for by the same set of metrics (and the weightings between these metrics), and are likely to need variation within the group. For example, income levels from different sources can vary greatly within a broad group such as Chemistry, Physics and Earth Sciences.

- There were concerns that minority disciplines (such as Astronomy) within the broad subject groups would be dominated by the larger ones.

- Many respondents argued that the REF should either operate at a more detailed subject level (for example, similar to RAE units of assessment, or at suitably modified

HESA cost centre level), or that it should operate at a broad subject level as proposed but also generate finer grained information (for example, at the field level at which bibliometric data is normalised).

- Many were concerned that the broad groups would limit the usefulness of both the REF for internal research management purposes and resource allocation, and the information provided to external stakeholders (such as prospective students, industry, and potential collaborative partners). A few stated that a quality profile at subject group level may well identify institutions where excellent research is taking place, but not the disciplines within the institution.

- There were concerns that smaller departments and pockets of excellence would not be visible, and this would impact on the reputation of individuals and departments.

- Many felt that broad groupings would constrain the expert panels (this is discussed further below).

- Some respondents were not concerned about the broad subject groups, arguing that the REF should have a single purpose of allocating QR funds.

18. Many respondents provided feedback on the location of particular disciplines on either side of the divide:

- There was consensus that Psychology should be on the science-based side, or that it should at least be tested in the pilot with a view to moving it. Some, however, noted that some areas within Psychology are not well suited to this.

- Many respondents argued that modes of research in Nursing and Allied Health Professions are more akin to social science research, and argued that these subjects should be assessed by light touch peer review informed by metrics. Metrics could include bibliometric indicators – although the limitations in citation data were also noted – as well as indicators better suited to practice-based research.

- A number of respondents discussed the location of Mathematics and Statistics, and most agreed that light touch peer review is more appropriate for these subjects. However, it was noted that there is much collaboration with science and Engineering, and that some important research in mathematics is embedded within other disciplines, leading to concern about mathematics and statistics being assessed on the other side of the divide. The mathematics community also wished to ensure that an appropriate form of light touch peer review would be applied, and suggested that this would differ from peer review appropriate for the arts and humanities.

- Some respondents noted that Pharmacy is highly multidisciplinary – involving health, biological and physical sciences as well as elements of social science – and were

concerned about its fit within the REF and about modes of assessment not well suited to the applied focus of much research.

19. A number of respondents commented on the way we had mapped RAE 2008 units of assessment to the proposed six subject groups:

- They noted that the groups are disparate, and queried in particular the coherence of the proposed Health Sciences and Subjects Allied to Health groups. Other examples of diversity include Biological Sciences, which spans disciplines such as forestry and molecular genetics.

- A number of suggestions were made to move particular disciplines into different groups, including a number within Health Sciences and Subjects Allied to Health, and Veterinary Science.

- Some argued that Computer Science should be separate from Engineering.

- Some noted that the links between subjects are complex and would cut across any broad grouping (for example, there are strong interfaces between Physics and Engineering, and between Chemistry and Biological Sciences).

- Some queried whether the disciplines within each group have similar publication and citation patterns, and suggested we should do further work to group disciplines on this basis, at least for the purposes of bibliometric assessment.

20. Many respondents commented in detail on the Engineering and Computer Science group and urged careful consideration of this:

- Two main limitations of bibliometric indicators were raised. Firstly, the coverage of citation data is relatively low, and important types of outputs other than journals are not covered (including conference proceedings, software and commercially confidential reports). Some felt it might be possible to address this by supplementing the available data. Secondly, citations by academic peers were not seen as an appropriate way of measuring quality for work that is focused on application and users. The engineering community in particular was concerned that emphasising bibliometric indicators would risk driving researchers towards more theoretical work and away from economically important user-driven or applied work.

- The engineering community urged that additional indicators would need to be developed to focus more on impact and the quality of applied research, and that relatively little weight should be given to bibliometric indicators. It was recognised that this would not be straightforward and there were doubts that it could be done quantitatively.

- It was felt that that in these subjects expert panels (including users) would need to apply judgments to the indicators, and that some qualitative assessment by experts would be desirable or necessary to supplement the indicators.

- The computer science community expressed some concern that it would be unduly dominated by engineering if grouped together, and noted the diverse cross-disciplinary nature of computer science research, which has important interfaces with mathematics and other subjects.

21. Respondents commented on a number of other discipline-specific issues, including:

- Several disciplines were mentioned in which publication and citation patterns may vary considerably across sub-disciplines, or between experimental and theoretical approaches. Respondents queried whether bibliometric techniques could be sensitive to this, or suggested that expert input would be needed to account for it.

- Some noted that particular disciplines tend to use specialist bibliographic databases (for example, for particle physics and for astrophysics) and urged that these should be considered.

- Some were concerned about our proposals to exclude self-citations for particular areas of research that tend to involves large-scale collaboration (such as particle physics).

- In addition to Engineering and Computer Science, a number of disciplines (particularly in the medical, health and related subjects) stressed the importance of capturing user-valued, applied and practice-based research, and raised related concerns about the lack of expert review.

- Areas of research that require collaboration across a range of disciplines, such as environmental and development sciences, raised particular concerns about research that crosses the divide between science-based and other subjects; and stressed the importance of capturing impact and user value.

Bibliometrics

22. We invited responses to the following questions:

> Consultation question 2a: Do you agree that bibliometric indicators produced on the basis that we propose can provide a robust quality indicator in the context of our framework?
>
> Consultation question 2b: Are there particular issues of significance needing to be resolved that we have not highlighted?

23. Many respondents provided general comments about the proposed use of bibliometric indicators and their suitability for the REF:

- A number of respondents supported some key features of the proposed approach, including basing the indicator on citation rates per paper (rather than using journal impact factors); normalising data in a way that is both sensitive to field-specific differences and benchmarking performance internationally; the intention to produce a quality profile rather than a single point score; and the potential for a less burdensome and more transparent system.

- It was widely commented that bibliometrics do not provide a direct measure of quality, but provide a proxy indicator of quality. Many also felt that bibliometric indicators focus only on intellectual influence among the academic community, and hence do not provide a rounded indicator of quality. In particular, they do not capture the quality of applied or practice based research or user-value, and many argued that bibliometrics should be used alongside other quality-related indicators.

- In general, respondents felt that further development and testing will be required before they could be convinced of the robustness of bibliometric indicators, and that the outcomes would need to be moderated and validated by expert panels in order to address a range of potential problems and secure the sector's confidence.

24. There was substantial discussion around the scope and coverage of the bibliometrics process, and a range of views were expressed:

- The majority of people who mentioned the issue felt that all research staff in relevant disciplines should be included in the process, rather than selected staff. This was typically justified on the grounds that it would reduce the burden of selection by institutions, and could also address equal opportunities issues.

- Some suggested that we should only include certain (automatically selected) papers from all staff. Some of those who implied that staff should be selected individually for inclusion, and thought this would be burdensome, had taken this implication from the consultation document as a given.

- Those who supported selection of staff generally thought that institutions should do the selecting (rather than an algorithmic approach). These were typically the larger research intensive universities.

- Many respondents discussed whether papers should be credited to the institution that employed the researcher at the time of publication, or whether papers should follow the author when they move institutions. Opinion was split, with about half arguing the case each way. A few suggested that both the institution where the author was employed, and the author's new institution should get some credit. There was little consensus over which would be the less burdensome approach. Issues

with people entering UK academia, either from industry or abroad, were mentioned in a few responses.

25. Many respondents commented on the source data that will be used:

- Many raised concerns about the Web of Science (WoS) coverage, especially with respect to conference proceedings and engineering and computer science. A few responses indicated that books, reports, software, and so on should be included in the database, and some gave examples of journals that they felt are wrongfully excluded. Several respondents were keen for us to explore other databases. A few commented on using a target expanded approach, but noted that it is time-consuming to implement.

- Our reliance on a single commercial data-supplier was raised as a general concern; some responses elaborated that the selection of journals to be taken into account would be out of our control.

- Several responses raised concerns about the data quality and requested further work to assess this.

- Some responses noted that open access forms of publication are likely to become more common in future, and that the REF system would need to be able to take account of them. A few thought that REF would have a negative effect on open access publication.

26. Many responses provided feedback on key technical issues:

- Many stated it would be important to consider citation windows carefully. The retrospective nature of bibliometrics was noted by many, and several commented that we would initially be re-assessing work already assessed by the 2008 RAE. It was noted by a few that the time lag could act as a damper on a dynamic research base. Some noted that the appropriate window could vary between disciplines, and suggested further work to look at how citation half-lives vary by discipline.

- There was concern about our proposal to exclude self-citations, especially in areas where papers with large numbers of authors are the norm (such as in particle physics, astrophysics, and genome work). It was noted that the exclusion of self-citations could act as a deterrent to collaboration, and further work was requested on this issue. Several responses did concede that the problem would be non-trivial to solve.

- A number queried how multiple authorship (especially between multiple institutions) would be handled.

- Many noted it would be vital to define appropriate fields for the purposes of normalisation. A number of respondents thought that the existing fields within the WoS database could be too broad and in some cases gave examples, such as particle physics. Some noted that the contents within each normalisation field should be statistically similar. The difficulties of appropriate normalisation for engineering and computer science, and for some disciplines at the edges of the science-based subjects were mentioned by some respondents. The difficulty of normalising interdisciplinary work and new and emerging areas was sometimes mentioned; what should they be normalised against?

- Some were concerned about the robustness of bibliometric indicators for smaller scale activity (noting that some institutions have relatively few researchers even at the broad subject group level). It was mentioned that these may be unusually susceptible to fluctuations, owing to the small amount of data making up their profile.

27. Many respondents discussed the potential impacts and consequences of using bibliometric indicators:

- The majority of institutions expressed a concern about the potential for game-playing and other unintended behavioural effects. It was noted by several that such effects are difficult to predict. Citation clubs were often mentioned, as was the prospect of researchers pursuing safe research. A number suggested that researchers would be encouraged to produce review papers and methods papers, as these tend to be more highly cited.[2] Many thought that researchers would be under pressure to publish less, and may be reluctant to publish work that, although worthwhile, would be little cited (it was noted that this could also have implications for publishing with research students and early career researchers). Many suggested that the behavioural effects of the REF should be monitored, perhaps with a role for the expert panels in this.

- The potential negative effects on early career researchers and the burden implications of the bibliometrics process were raised by many (these issues are discussed further below).

- A number of responses commented that institutional repositories would be almost essential in supporting the process. The burden of this was often commented on. Some responses suggested that assistance (and funding) be given from JISC to help set these up. A few suggested a centralised (national) repository.

[2] It should be noted that there are techniques for normalising citations by type of paper to account for higher citation rates among review papers.

Light touch peer review

28. We asked:

> Consultation question 3a: What are the key issues that we should consider in developing light touch peer review for the non science-based disciplines?
>
> Consultation question 3b: What are the main options for the form and conduct of this review?

29. Respondents tended to discuss this in the context of their desire for a more unified system across all disciplines, and many raised concerns about the proposed timetable for developing the light touch peer review process only after determining the approach for the science-based disciplines. They urged that the timetables should be more closely aligned, for a number of reasons:

- To enable a more cohesive and holistic framework across all disciplines to be developed.
- To enable effective institutional planning given the long lead times for peer review and the operational changes required to manage a multi-method assessment system.
- To develop appropriate metrics to inform light touch peer review alongside, rather than as a result of, bibliometrics.
- To explore ways of capturing impact and user-value across all disciplines and integrating this into the REF as a whole.
- To consider the roles of panels across the REF at the same stage of the development process.
- To give more time to reflect upon the 2008 RAE process and the opportunities to reduce burden through REF, and to engage the help of the 2008 RAE panel members (once they have completed the exercise) in developing light touch peer review.

30. Many respondents discussed the scope for reducing the burden associated with peer review. In general, respondents doubted that significant reductions in burden could be achieved without compromising the rigour of the exercise and the confidence of the sector. Nevertheless, many respondents discussed the possible options for reducing burden:

- Respondents typically related the burden on HEIs to collating information and preparing submissions; some felt that this is where the development of a light-touch process should focus its efforts.

- Many discussed the potential to reduce the number of outputs submitted. Respondents noted that this would risk a perceived reduction in the rigour of the assessment and would not necessarily equate to a reduction in burden for the submitting institutions (although it could reduce the burden on panels). Some felt that this could be more burdensome in comparison to the current RAE arrangements, by increasing the stakes and pressure on institutions' processes for selecting outputs.

- A sampling approach to assessing outputs was suggested by some respondents as an area for further exploration.

- Some noted that the importance of different types of output across the disciplines should be considered, if fewer outputs are to be submitted or assessed.

- There was consensus that a possible reduction in the number of panels should be considered carefully. Many respondents were concerned that this would lead to the loss of the necessary levels and coverage of expertise required to make informed judgements regarding the quality of research outputs.

- Many raised the potential to omit or change the current textual commentary element of the submission (the RA5a). While many respondents believed that this narrative element of the submission was an important and useful way to communicate their research context, they generally felt it is unduly burdensome to prepare. Some respondents thought the narrative element should be retained (across all disciplines).

31. Many discussed the use of metrics to inform peer review, and generally favoured the use of robust and subject-sensitive metrics. The following main points were raised:

- There was recognition that bibliometrics could play a significant role in certain disciplines, and it was felt that the potential for this is likely to increase over time.

- Many suggested that the wider impact of research should be taken into account in light touch peer review, although several raised concerns about this (for example, potential risks to quality) or suggested that other policy and funding levers are better placed to address wider impact.

- Some wished to include indicators of esteem as an important way of informing the health of disciplines. Some stated that capturing involvement in peer review and other activities was crucial as proof of contribution to the academic and public good.

- Many urged that the development of metrics to inform peer review be undertaken after further consultation and as part of the wider development of bibliometrics in the science-based disciplines. The development of these metrics was identified by many as a high-priority area of work for HEFCE as we develop the framework.

User value, impact and other indicators

32. We asked:

> Consultation question 4: Is there additional quantitative information that we should use in the assessment and funding framework to capture user value or the quality of applied research, or other key aspects of research excellence? Please be specific in terms of what the information is, what essential element of research it casts light on, how it may be found or collected, and where and how it might be used within the framework.

33. Respondents raised several general issues about the use of metrics in the REF:

- Many were uneasy about an over-reliance on bibliometrics and most felt that a range of indicators should be used in each discipline group.
- Respondents generally said or implied that indicators would need to be subject-specific, and that expert panels should advise on appropriate metrics for their subject. Some suggested consulting the RAE panels on what information they found most useful in forming judgements.
- A few responses suggested that a wide range of metrics should be tested for each discipline and selected on the basis of how well different combinations of metrics correlate with RAE results.
- There was consensus that metrics will need to be moderated or interpreted by expert panels, and they may need to include qualitative as well as quantitative elements.
- Some respondents urged caution about the potential complexity and burden of developing a raft of (subject-specific) metrics.

34. Much of the discussion about metrics focused on the issue of user value and impact, and there were differing views about the importance of capturing this within the REF:

- Most argued that it is important for the REF to recognise and reward research that has a positive economic and social impact. Many saw this as essential, in line with government policy and vital to promote higher education's contribution to the economy and society.
- In subjects with a significant applied, user-focused or practice-based element (especially engineering, but also medicine, health sciences and computer science), there were strong concerns that if the REF focuses narrowly on academic excellence, this would drive research activity towards theoretical work and away from applied and user-focused work; and that this would be damaging to these disciplines and to the economy and society.

- Many stressed the importance of capturing social, cultural, policy and wider impacts, and not focusing narrowly on economic impact, and that this should be a feature of the REF across all disciplines as appropriate.

- A significant minority (including the Russell Group and a number of institutions and academies) argued that the REF should focus solely on the quality of research (both fundamental and applied), and that interaction with users and user-value is promoted through other funding streams and policies, including the Higher Education Innovation Fund (HEIF) and the Research Councils. However some of these respondents did acknowledge that user value is an important feature in certain subject areas.

- Some were generally cautious about capturing impact because of the complexity of assessing it, or were concerned that a narrow or short term view of impact could present risks to fundamental research.

- A few suggested reconsidering the relationship or balance of funding between QR and HEIF, or possibly merging them.

35. While most respondents were keen to capture impact or user value, they generally recognised that this would be challenging for a number of reasons: there can be long time lags before real impact is known; problems of causality; ways of measuring it differ between subjects; available data is limited; and many kinds of impact are not amenable to quantification. Respondents had differing views on how to try to incorporate impact and user value in the REF:

- Some argued that the simplest proxy for impact would be to use income from user groups (industry, charities and government departments), though others highlighted shortcomings of this approach.

- Some stated that no adequate metrics are currently available and requested that we should develop new metrics. It was suggested that this could be led by an expert user group, but would take time.

- Some suggested we should review any relevant indicators used in the 2008 RAE as a starting point.

- Some argued that quantifiable measures of impact would not be informative, and that impact can only be assessed through expert review (which could be informed by qualitative as well as quantitative information).

- Some stressed that different approaches would need to be determined at subject level, to reflect different emphases on economic, social, cultural, health, policy and

other impacts, the appropriate means of assessing them, and the different timeframes in which different impacts become evident.

36. Respondents suggested a wide range of impact-related measures to consider, though many noted they would need to vary by discipline and are not necessarily quantifiable:

- Suggestions for economics related indicators included:
  - patents lodged and volume of intellectual property income generated (though it was noted that using these could rapidly lead to distortions of behaviour)
  - repeat research income from industry
  - licenses granted to companies by the research group (to pick up on research which is actively exploited)
  - leverage against QR investment
  - engagement with small to medium-sized enterprises (possibly through the take- up of R&D tax credits)
  - measures which recognise near-to-market and developmental research
  - number of academic personnel with industrial consultancy agreements
  - joint authorship of publications with business or other user partners
  - basic research which advances the field of knowledge of the sponsor (that may not yet have obvious application)
  - measures of spin-out companies (which have sustained their existence for a period, and taking into account jobs created and investment attracted).

- Measures of impact on health practice included: impact on health service guidelines and international treatment guidelines; evidence of improvements in patient outcomes; reducing costs in the acute care sector and community; service user involvement.

- Suggestions for measuring impact on policy development included citation of research in various government and public sector documents and proceedings, and formal advisory roles.

- Some suggested that indicators of journal usage (currently being developed) would provide a useful indicator of usefulness, to supplement citation indictors.

37. A number of respondents discussed the use of income and postgraduate research (PGR) student data. Although there were general doubts that such input metrics are indicative of quality, there was consensus that they should nevertheless play a role within the REF. A number of specific points were raised:

- Some noted income levels depend on cost rather than quality, and may vary greatly within subject groups and disciplines depending, for example, on the nature of research equipment required in different fields.

- Some commented that the volume of PGR students does not reflect quality, and some suggested that completions would be more informative.

- Some argued that the REF should take account of all sources of income, while some suggested we should treat them differently (for example, peer reviewed sources focus more on quality; while income from research users reflects more on impact). Some urged caution about the implications for Research Council application processes, if the REF uses Research Council income data.

- Some noted that medical research income for employed clinical academics can come through the NHS Trusts associated with medical schools, and may not be captured in the available data.

38. A number of respondents discussed or suggested other types of indicators:

- Some suggested a sustainability indicator, looking at the profile of staff and PGR students. Environment was also mentioned by some as important for sustaining the research base.

- Some suggested measures of engagement with the public and user groups (such as employers and policy makers).

- A number of responses mentioned esteem, and felt it would be important for the REF to encourage activity such as editorships.

39. Many respondents welcomed the intention to rely more on HESA data, although a number pointed to potential difficulties with mapping cost centres to REF subjects; differences in the way activity is attributed to cost centres; and the need to allow sufficient time for improvements in data quality.

Expert panels

40. We asked:

> Consultation question 5: Are our proposals for the role of expert panels workable within the framework? Are there other key issues on which we might take their advice?

41. Many respondents discussed the role of expert panels:

- Most felt that bibliometrics (and other metrics) would not be sufficiently free from distortion, well understood or mature to be used without some form of moderation by expert panels. This was described in different ways, such as reviewing the outcomes for anomalies, and validating the outcomes. Some thought this would be especially important initially, until there is greater confidence in the system. Some raised the

question of what additional information the expert panels might need in order to moderate the metrics, though there was little beyond this of how in practice the panels could effectively moderate.

- A number suggested that expert panels should also play a role in addressing the limitations or potential problems with bibliometrics, by advising on significant issues (such as non-journal outputs, fields with lower data coverage, problems with early career researchers or interdisciplinary research, and applied research) and addressing them through additional data and/or elements of peer review.

- There was general agreement that panels should have an important role in determining suitable metrics for their subject and determining the weightings between them.

- A number of respondents argued that a substantial element of expert judgment was needed in all disciplines, because of limitations in the available metrics and the complex range of factors that should be considered.

- A few respondents stated explicitly that experts should not be able to moderate or overrule metrics.

- Some felt that expert panels would need some qualitative or judgment role in assessing impact in particular.

- Many suggested a role for panels in monitoring their community's behaviour and potential adverse impacts on publication behaviours, for example.

- Some respondents urged caution about not allowing the expert panels' roles to expand over time, leading to increased burden.

42. Many also commented on the subject breadth of panels and their membership:

- There was general concern that panels for the six proposed subject groups in the science-based disciplines would be too broad to contain specialist discipline-specific knowledge within each subject group. Respondents felt this would limit panels' understanding of the differences between disciplines within their subject group, and would constrain their roles in interpreting data and understanding behaviours. Many suggested either narrower panels or using sub-panels (perhaps initially and then phasing them out), or at least large membership of the broad panels.

- Many felt that panels should include research users. Some queried whether panels would need bibliometric expertise. Some suggested including international members. A few suggested panels should include interdisciplinary researchers.

- A number raised concerns about the implications of having very broad panels in the science-based disciplines and somewhat more narrow panels for the other disciplines. A number of respondents commented that panels in the arts, humanities and social sciences could potentially be somewhat broader than the RAE units of assessment, but that the scope for this is limited.

Burden and implications for institutions

43. We asked:

> Consultation question 6: Are there significant implications for the burden on the sector of implementing our new framework that we have not identified? What more can we do to minimise the burden as we introduce the new arrangements?

44. Overall the majority of responses welcomed the intention to reduce burden compared with the current RAE, though many pointed to a tension between reducing burden and maintaining the rigour and integrity of the assessment process. Discussion focused on the burden on institutions, though it was noted that the burden on assessment panels should also be considered. Many respondents felt there was relatively limited scope to significantly reduce the burden on institutions without posing a risk to rigour, and felt it was important that the REF must command the confidence of the community.

45. There was general consensus that burden in the short term (during the transitional phase) would not be reduced, and a number thought it would significantly increase, for the following reasons:

- Many respondents considered that the timescale for the introduction of the REF was in itself a major source of short-term burden. Changes to institutional project management structures would be necessary to cope with the two assessment methodologies proposed in the REF, which would require time to design, implement and embed. This may require additional resource or retraining for existing staff. Higher education institutions voiced a concern that there could be a transfer of activity from research offices to more central services such as libraries to cope with the demands of a bibliometrics-based system in the sciences, including a perceived requirement for bibliometricians.

- Many felt that considerable effort would be needed for institutions to develop the systems for managing research publication records that would be necessary to operate the bibliometrics process. It was clear from the responses that HEIs are currently at varying stages in developing such systems, and many were concerned that the short timescale for introducing bibliometrics would create additional pressures. A number felt that without clear and early guidance from HEFCE there is a risk that systems and institutional repositories currently being developed may not meet the requirements of REF.

- A number of respondents felt that the initial bibliometrics process would be assessing research that had already been assessed by the 2008 RAE, and to this extent represented unnecessary additional burden in the short term.

46. While there were frequent concerns about the transitional burden, many respondents generally accepted that there is potential for the burden of the REF to reduce over time, and in steady state to be lower than the RAE, although a number of caveats were raised, and some felt this is unlikely. The following key points were raised:

- The operational difficulties of managing two distinct assessment processes was highlighted (for example, the difficulties of assigning staff to either side of the boundary and the internal management issues surrounding this, including differential reward mechanisms and appeal processes). It was felt that these difficulties would be exacerbated by running each system to a separate timetable.

- Many respondents urged us to consider an automated staff selection process for the bibliometrics-assessed subjects in order to reduce burden. A number of these felt that if institutions continue to select staff, the burden would be similar to that of the RAE. However, a significant minority of responses (including a number of research-intensive institutions in particular) felt it important that institutions should continue to select staff to be assessed in all subjects.

- There was much discussion about the potential burden of data-checking and verification required to operate the bibliometrics process in steady state. Many acknowledged that the level of burden associated with this would depend on several aspects of the system which are not yet known:
  - whether staff are to be selected
  - how frequently the process is repeated
  - whether papers are credited to the institution (this was seen as less burdensome than papers following the author)
  - how far the process would provide useful information to institutions (if this is limited then institutions may need to undertake considerable work to generate such information for themselves).

- It was generally considered that there is limited scope for reducing the burden of the light touch peer review process on institutions (although scaling down the number of outputs submitted could lighten the load on panels).

47. In addition to considering burden, many responses discussed the implications of the REF for institutional research management. It was widely acknowledged that the RAE has been used by many institutions to inform internal research management and resource allocation, and many institutions were concerned about moving to much broader subject groups, at least for the science-based disciplines:

- Many were concerned by the prospect of not being able to identify quality at discipline level, to inform research management and resource allocation as well as provide public information. They either suggested that assessment should take place at a more detailed level than the six broad subject groups (for this and other reasons), or requested that if the REF operates at broad subject level, data should also be made available at a more detailed level.

- A number of institutions also wished to be able to access the underlying data, both in order to produce information for internal use and to replicate and understand exactly how the REF profiles had been arrived at. It was suggested that either HEFCE should make the underlying data available, or there should be a centrally-funded license for institutional access to the data (although the latter option raised questions about the amount of cleaning of data that would be needed to replicate the REF outcomes).

- Some cautioned that, although institutions are likely to break down the REF profiles, bibliometric data may not be robust below broad subject level, raising concern about potential misuse of the data.

- Some institutions considered that the lack of results at discipline level would be an inevitable part of the transition to a metrics-based system, and were prepared to develop their own internal systems for research management, without relying on the REF. They argued that the REF should have a clear and single purpose of allocating QR, in order to minimise burden.

Equal opportunities

48. We asked:

> Consultation question 7: Do you consider that the proposals in this document are likely to have any negative impact on equal opportunities? What issues will we need to pay particular attention to?

49. Many respondents felt that for light touch peer review, the issues were broadly similar to those exposed and catered for in the current RAE, and they felt that the promotion of equal opportunities through the requirement for institutional codes of practice should not be lost in the transition to the REF. However, many responses identified concerns with the bibliometrics process:

- Many felt that bibliometrics could disadvantage early career researchers in favour of more established researchers because of the retrospective nature of this type of assessment, especially if the citation window is relatively long.

- Many also felt that those who take career breaks or work part-time could be disadvantaged, as they are likely to have fewer outputs and citations, and we were

urged to consider how citation windows and any thresholds that may apply would affect this.

- Some queried whether there could be inherent bias within the underlying data, for example are females disadvantaged in the publication process?

- There were a number of suggestions for alleviating some of the potential difficulties, including: a universal or automated approach to including staff (rather than selection of staff by institutions); requiring institutional codes of practice especially if they select staff; retaining a narrative element to explain researchers' personal circumstances; developing a 'sustainability' indicator to show the profile of staff.

Other issues

50. We asked:

> Consultation question 8: Do you have any other comments about our proposals, which are not covered by the above questions?

51. Many responses stated it would be important to thoroughly pilot the REF, and made suggestions for issues that should be tested:

- Scope and coverage issues: many suggested we should test a non-selective approach to including staff.

- Data issues and requirements: a number felt we should assess the robustness of the underlying data and test the verification processes. Some also suggested we should test means of supplementing the WoS database.

- Implications for interdisciplinary research, within each broad subject group and across the divide, and for new and emerging areas of research.

- Institutional implications: including costs, the information that can be made available to institutions, and the implications for small pockets of excellent research.

- Subject groupings: including boundary issues and refining the proposed subject groups.

- Technical issues: including citation windows, handling multi-authorship and self-citation.

- Potential impacts: including incentives on publication and other behaviours, and equal opportunities issues.

- Expert panels: the interaction between metrics and expert input, and moderation of bibliometric indicators.

- Validation: testing outcomes against the 2008 RAE (those who mentioned this recognised that an exact match is not to be expected, but would still wish to be assured by a comparison between the two).

- Many suggested that the pilot should include all the 'intermediate' disciplines (such as geography and sports science) and possibly also social sciences, in order to illuminate issues about boundaries and interdisciplinarity, as well as explore the potential for bibliometric indicators to inform light touch peer review in some disciplines.

- Some suggested an early priority is to establish a steering committee for the pilot group.

52. There was widespread concern that the current timetable is too tight, and many requested that we extend it by at least 12 months:

- Many stressed the importance of thoroughly piloting the system, with sufficient time to refine the proposals (and consult further) before moving to implementation. There did not appear to be time for this in the current timetable.

- Some were concerned that HEIs are given sufficient time to prepare; in the current timetable there appeared to be very little notice of the requirements for the bibliometrics exercise.

- It was generally felt that ensuring the robustness of the new system was more important than meeting the current timetable, especially given that the results of the 2008 RAE will be able to inform funding in the interim.

- Many felt that the outcomes of the pilot should be evaluated against the 2008 RAE, and that we should draw on the expertise of the 2008 RAE panels before we finalise the system; this could only be done during 2009 at the earliest.

- Many wished to see the timetables for the science-based and other disciplines to be more aligned, to enable more cohesive and holistic development of the REF, and requested that we bring forward the development of light touch peer review as well as extending the timetable for the science-based approach to enable this.

53. Some other points were made about implementation of the REF:

- Many welcomed the opportunity to input and urged continuing consultation, especially after the pilot phase.

- Some urged that changes to funding should be managed and moderated with stability, to enable institutional planning.

- A number queried how often the REF will be repeated; most who raised this suggested it should not be too frequent given the cost and likely stability of outcomes; and that we should synchronise the metrics and review exercises.

Annex A
Organisations responding to the consultation

Higher education institutions

University of Aberdeen
University of Abertay Dundee
Universities of Aberystwyth and Bangor
Anglia Ruskin University
Aston University
University of Bath
University of Bedfordshire
Birkbeck College
University of Birmingham
Bishop Grosseteste University College Lincoln
University of Bolton
Arts Institute at Bournemouth
Bournemouth University
University of Bradford
University of Brighton
University of Bristol
Brunel University
Buckinghamshire New University
University of Cambridge
Institute of Cancer Research
Canterbury Christ Church University
Cardiff University
University of Central Lancashire
Central School of Speech and Drama
University of Chester
City University London
Coventry University
Cranfield University
University of Cumbria
Dartington College of Arts
De Montfort University
University of Derby
University of Dundee
Durham University
University of East Anglia
University of East London
Edge Hill University
University of Edinburgh
Edinburgh College of Art
Institute of Education
University of Essex
University of Exeter

University College Falmouth
University of Glamorgan
University of Glasgow
Glasgow Caledonian University
Glasgow School of Art
University of Gloucestershire
Goldsmiths College
University of Greenwich
Guildhall School of Music and Drama
Harper Adams University College
Heriot-Watt University
University of Hertfordshire
University of Huddersfield
University of Hull
Imperial College London
Keele University
University of Kent
Kings College London
Kingston University
Lancaster University
University of Leeds
Leeds Metropolitan University
Leeds Trinity and All Saints
University of Leicester
University of Lincoln
University of Liverpool
Liverpool John Moores University
University College London
University of the Arts London
London Business School
London School of Economics and Political Science
London School of Hygiene and Tropical Medicine
London Metropolitan University
London South Bank University
Loughborough University
University of Manchester
Manchester Metropolitan University
Napier University
Newcastle University
Newman University College
North East Wales Institute of Higher Education
University of Northampton
University of Northumbria
University of Nottingham
The Nottingham Trent University
Open University
School of Oriental and African Studies
University of Oxford

Oxford Brookes University
School of Pharmacy
University of Plymouth
University College Plymouth St Mark and St John
University of Portsmouth
Queen Margaret University
Queen Mary, University of London
Queens University Belfast
University of Reading
The Robert Gordon University
Roehampton University
Royal Academy of Music
Royal Agricultural College
Royal College of Art
Royal Holloway, University of London
Royal Northern College of Music
Royal Scottish Academy of Music and Drama
University of St Andrews
St Georges Hospital Medical School
St Marys University College
University of Salford
University of Sheffield
Sheffield Hallam University
University of Southampton
Southampton Solent University
Staffordshire University
University of Stirling
University of Strathclyde
University of Sunderland
University of Surrey
University of Sussex
University of Swansea
University of Teesside
Thames Valley University
UHI Millennium Institute
University of Ulster
University of Wales Institute Cardiff (UWIC)
University of Wales Lampeter
University of Wales Newport
University of Warwick
University of the West of England Bristol
University of West of Scotland
University of Westminster
University of Winchester
University of Wolverhampton
University of Worcester
University of York
York St John University

Other organisations

1994 Group
Academic General Practice and Primary Care
Academy of Marketing (UK)
Academy of Medical Educators
Academy of Medical Sciences
Academy of Pharmaceutical Sciences of Great Britain
Academy of Social Sciences
Architectural Humanities Research Association
Association for Learning Technology
Association for Political Thought
Association for the Study of Modern and Contemporary France
Association for Tourism in Higher Education
Association of British Pharmaceutical Industry
Association of Business Schools
Association of Heads of Psychology Departments
Association of Medical Research Charities
Association of Research Managers and Administrators
AURIL
bioProcess UK
Biosciences Federation
Board of Celtic Studies (Scotland)
British Academy
British Academy of Management
British Association for Applied Linguistics
British Association for Counselling and Psychotherapy
British Association/College of Occupational Therapists
British Classification Society
British Computer Society
British Dental Association
British Heart Foundation
British Philosophical Association
British Society for Dental Research
British Sociological Association/Heads and Professors of Sociology
Campaign for Science and Engineering
Cancer Research UK
Chartered Society of Physiotherapy
Chemistry Innovation Knowledge Transfer Network
Committee of Heads of University Law Schools
Committee of Professors of Statistics UK and Ireland
Confederation of British Industry
Confederation of British Industry Scotland
Confederation of British Industry Wales
Conference of Heads of University Departments of Economics
Conference of Professors of Accounting and Finance
Conference of University Teachers of German in the UK and Ireland
Conservatoires UK
Council for College and University English

Council for Higher Education in Art and Design
Council for Hospitality Management Education
Council for Industry and Higher Education
Council for the Mathematical Sciences
Council of Deans of Health
Council of Heads and Deans of Dental Schools
Council of Professors and Heads of Computing
Council of University Classical Departments
Council of University Heads of Pharmacy
COUNTER Online Metrics and UK Serials Group
Consortium of University Research Libraries/Society of College, National and University Libraries
Department for Food and Rural Affairs
Economic History Society
EEF
Engineering Professors Council
Environment Research Funders Forum
Equality Challenge Unit
Geological Society of London
GlaxoSmithKline R&D
GuildHE
Guttridge Ltd
Heads of Chemistry UK
Heads of Departments of Mathematics
Heads of Health Economics Units
Heads of University Biological Sciences
Heads of University Centres for Biomedical Science
Higher Education Academy Psychology Network
Higher Education Wales
History of Education Society
Institute of Physics
JIBS User Group (JISC assisted Bibliographic data Services)
John Pethullis Chair of IMechE BMHC
Joint Committee for Psychology in Higher Education
Joint Universities Council Social Work Education Committee
Leisure Studies Association
Linguistics Association of Great Britain
Marie Curie Cancer Care
Media Communications and Cultural Studies Association
Medical Schools Council
Micropalaeontological Society
Million+
Modern Universities Research Group
National Association for Music in Higher Education
National Association of Writers in Education Higher Education Network
National Endowment for Science Technology and the Arts
National Union of Students
Newton's Apple

Palaeontological Society
Philosophy of Education Society of Great Britain
Professors of Midwifery UK
RAE2008 Panel E
RAE2008 Sub Panel Psychology
Research and Development Depts, UK Health Departments
Research Councils UK
Research Forum for Allied Health Professions
Research Information Network
Royal Academy of Engineers
Royal Astronomical Society
Royal College of Nursing Research Society
Royal Geographical Society with Institute of British Geographers
Royal Historical Society
Royal Pharmaceutical Society of Great Britain
Royal Society
Royal Society of Chemistry
Russell Group
Scots Philosophical Club
Scottish Universities Physics Alliance
Social Policy Association
Society for Old Testament Study
Society for Research into Higher Education
Society of Legal Scholars
Solids Handling and Processing Association
Standing Committee for Archaeology
Standing Conference on Dance in Higher Education
Technology Strategy Board
The English Association
UK Collaborative on Development Science
UK Computing Research Committee
UK Council for Graduate Education
UK Council of Area Studies Associations
UK Joint University Public Administration Committee
UK Political Studies Association
UK Professors of Midwifery (maternal and perinatal health)
Unico
Universities and Colleges Information Systems Association
Universities Council for the Education of Teachers
Universities Scotland
Universities UK
University Alliance
University and College Union
University and College Union Scotland
University Forum for Human Resource Development
Wellcome Trust

There were a further 53 responses from academic departments and individuals.