

## Annex A

### Consultation questions and response form

1. Responses to the consultation should be made by completing the form below, and returning it by e-mail by **midday on Wednesday 16 December 2009**.
- 2.
3. All responses should be e-mailed to [ref@hefce.ac.uk](mailto:ref@hefce.ac.uk). **In addition:**
  - a. Responses from institutions in Scotland should be **copied to** Pauline Jones, Scottish Funding Council, e-mail [pjones@sfc.ac.uk](mailto:pjones@sfc.ac.uk).
  - b. Responses from institutions in Wales should be **copied to** Linda Tiller, Higher Education Funding Council for Wales, e-mail [linda.tiller@hefcw.ac.uk](mailto:linda.tiller@hefcw.ac.uk).
  - c. Responses from institutions in Northern Ireland should be **copied to** the Department for Employment and Learning, e-mail [research.branch@delni.gov.uk](mailto:research.branch@delni.gov.uk).
  - d.
4. We will publish an analysis of responses to the consultation. Additionally, all responses may be disclosed on request, under the terms of the Freedom of Information Act. The Act gives a public right of access to any information held by a public authority, in this case HEFCE. This includes information provided in response to a consultation. We have a responsibility to decide whether any responses, including information about your identity, should be made public or treated as confidential. We can refuse to disclose information only in exceptional circumstances. This means responses to this consultation are unlikely to be treated as confidential except in very particular circumstances. Further information about the Act is available at [www.informationcommissioner.gov.uk](http://www.informationcommissioner.gov.uk). Equivalent legislation exists in Scotland.
- 5.

#### Respondent's details

<b>Are you responding:</b> (Delete one)	On behalf of an organisation
<b>Name of responding organisation/individual</b>	United Kingdom Computing Research Committee
<b>Type of organisation</b> (Delete those that are not applicable)	The UKCRC is an expert panel of the British Computer Society, the Council of Professors and Heads of Computing, and the Institution of Engineering and Technology, for computing research in the UK. This report was prepared by a working group whose members include Professors Alan Bundy (Convenor), Nick Jennings, Cliff Jones, Ross King, Ursula Martin, Tom Rodden and Morris Sloman. Six of these were members of the RAE panel in 2008, of whom three were also members in 2001.
<b>Contact name</b>	Alan Bundy
<b>Position within organisation</b>	Convenor of REF Working Group

<b>Contact phone number</b>	0131 650 2716
<b>Contact e-mail address</b>	A.Bundy@ed.ac.uk

## Consultation questions

(Boxes for responses can be expanded to the desired length.)

**Consultation question 1:** Do you agree with the proposed key features of the REF? If not, explain why.

In general yes.

While having no issues with the shift in definition of research (para 24), within the Computing Community there are concerns about

\* The scope of what is considered within the impact element of the assessment including the statement in 27b that “impact will be weighted more highly than environment” and the current weighting of 25% in the overall profile. We believe this value is too high.

\* The suggested timescale (27k) given the need to pilot a new assessment process for impact and develop systems with HEIs to manage the process. We believe it would make sense to delay the timescales by at least one year.

**Consultation question 2:** What comments do you have on the proposed approach to assessing outputs? If you disagree with any of these proposals please explain why.

Comments are especially welcomed on the following proposals:

- that institutions should select research staff and outputs to be assessed
- for the categories of staff eligible for selection, and how they are defined
- for encouraging institutions to submit – and for assessing – all types of high-quality research outputs including applied and translational research
- for the use of citation information to inform the review of outputs in appropriate UOAs (including the range of appropriate UOAs, the type of citation information that should be provided to panels as outlined in Annex C, and the flexibility panels should have in using the information)

and on the following options:

- whether there should be a maximum of three or four outputs submitted per researcher
- whether certain types of output should be ‘double weighted’ and if so, how these could be defined.

We feel that it is important that institutions should be responsible for selection of both staff to be submitted and the outputs that are assessed as they are best placed to make informed judgements

A person submitted as category A staff should have a contract of at least 25% FTE and actually have spent 25% of their time at the institution, not elsewhere, since the start of their contract. The number of outputs which a person can submit should be proportional to their FTE contract.

We welcome the simplification on criteria and would encourage these to be as clearly defined as possible. We would welcome an elaboration of the criteria provided in 33b. In particular we feel that there are three other aspects of eligibility that need to be addressed:

1. Whether externally directly funded researchers are eligible (e.g. those supported by research councils, the EU or by industry). If so, what is the threshold for eligibility? For consistency with previous RAEs 'independent researchers' should be eligible. The wording "academic staff" in para 33a might imply that only non-RAs are eligible. Given the move from 'tenure' in one direction and 'fixed term contracts' in the other towards 'open-ended contracts' for both, any such distinction might be hard to define legally.
2. We are also concerned that the wording of 33b "funded by third party bodies in partnership with the HEI" may be seen to include only staff where a significant proportion of support is provided by the HEI. This may have a negative impact on a number of researchers, including those holding externally funded fellowships, for example from the Research Councils.
3. Para 116a gives the criterion " Selected staff in post on a census date in autumn 2012". In practice, in RAE 2008, additional ad hoc conditions were sometimes imposed post hoc. Under the REF a comprehensive set of rules should be announced in advance.

Para 122b suggests that the eligibility criteria will be simple. We doubt that it will be possible to combine simplicity with clarity in this complex area.

A broad range of outputs was submitted to CS & I for RAE 2008. We would encourage sub-panel criteria and working methods to make the range of outputs clear, with examples of the type of outputs which can be submitted.

We understand the desire to reflect the efforts involved in outputs in para 37 and support the idea that a substantial research monograph, e.g. a book of > 250 pages, can count as 2 outputs. However, it is critical that the criteria make it clear what constitutes a substantial research output. This is particularly critical given the diversity of artefacts that may be submitted as research outputs (e.g. software systems, some of which are the collaborative endeavours of large teams).

We would welcome some clarification on the relation between the term "actual or potential use beyond the academic sphere" in para 39 and the criteria on impact. Given the experience of RAE 2008, we would wish to encourage that the 'significance' statements suggested in para 40 be made available for all submitted outputs, and that they allow authors to state the full significance of the work, including that within the academic sphere.

Para 43(e)(ii) suggests that “ Where panels decide to review samples of the outputs, robust sampling techniques should be used and sufficient samples of outputs should be reviewed for each submission.” We are not convinced that sampling of such complex data will produce an accurate assessment. We suggest that HEFCE seeks expert statistical advice before embarking on this practice.

We welcome the use of citation information if available to *inform* the assessment of outputs. The computing community spans a broad range of research traditions and methods. In particular, we feel that it is *essential* for the CS&I sub-panel to have access to Google Scholar citations. Google Scholar has emerged as a key resource in CS and IT internationally and is routinely used as one of the assessment criteria in tenure and promotion cases. In the 2008 RAE, 63% of Panel 23 CS&I outputs had zero WoS citations or could not be found in WoS. For outputs that had citations, Google Scholar correlated better with the panel ratings than WoS. Google Scholar harvests citations from industrial reports, as well as PhD theses, technical reports, books, etc, whereas WoS and Scopus do not, so it gives a fuller and fairer picture. Even if Google Scholar is not provided for all other panels, special arrangements are needed for CS & I sub-panel. We would be very concerned if the final framework inhibited its use to inform judgements.

Para 45b suggests that two citation databases will be procured. Given the special difficulties of full coverage in the CS&I area, we suggest that our community be fully engaged in this procurement exercise. If a database suitable for CS&I cannot be identified, then CS&I should join those panels *not* using citation data, since using a database with poor coverage would be damaging.

It is essential for panel criteria to make clear that outputs without citations may, nevertheless, be of high quality. Citation counts are just one element of an overall, judgement-based assessment.

Requesting 3 outputs per person would reduce the assessment burden by 25%, but would reduce the scope of the work being assessed and is likely to lead to 'grade profile inflation', as people can be more selective in submission of outputs. We consider 4 outputs the preferable compromise, provided the number of assessors assigned to a sub-panel is proportional to the number of people submitting. Assuming that the CS&I sub-panel continue previous practice by assigning each output to 3 assessors, the workload for each assessor should be no more than about 800-900 outputs.

The proposal to “revise the descriptor for four star [output] quality” in para 41 causes concern and may be an issue in practice. No justification for this revision is given. If the existing descriptor is not broken then it does not need to be fixed. The danger in the proposal is that it will result in grade deflation, that can be misrepresented as a decrease in UK research quality, which in turn could threaten the funding of UK research.

We are concerned about the amount of work for HEIs in checking citation data for all their submitted outputs, as suggested in para 48, particularly given the nature of computing citations and the timescales involved in undertaking this check. We would be particularly concerned if this checking delayed the time available for expert panels to examine outputs. There is also a risk if HEIs are required to identify papers to be submitted in advance of the cut-off date, as this will further truncate institutional time available for preparation, and restrict the research that is able to be submitted in a fast moving discipline.

**Consultation question 3:** What comments do you have on the proposed approach to assessing impact? If you disagree with any of these proposals please explain why.

Comments are especially welcomed on the following:

- how we propose to address the key challenges of time lags and attribution
- the type of evidence to be submitted, in the form of case studies and an impact statement supported by indicators (including comments on the initial template for case studies and menu of indicators at Annex D)
- the criteria for assessing impact and the definition of levels for the impact sub-profile
- the role of research users in assessing impact.

Impact will be difficult to assess and there is no past experience to draw upon. Impact can sometime take 10-15 years and may span multiple assessment periods. For example a spin-out company may have been started from research within the same or a previous assessment period, but considerable expansion and take-up of the research may only happen 8 or 9 years after the company was set up. Thus the same basic research or developments of it could be submitted as case studies in 2 assessment periods. This should not be precluded by the regulations.

In some cases a research group may refine and apply fundamental work done elsewhere in the UK or overseas that leads to local impact. For example, fundamental work on Binary Decision Diagrams for Model checking came from Carnegie Mellon University in USA, but European universities helped European industry to adapt and exploit new methods for verification in the 1980s. Under the current definition of impact, the group doing the fundamental work cannot claim impact as it was used by academics. We, therefore, consider it essential to broaden the definition of impact to include academic and scientific case studies as well.

The variation in time to impact is mirrored by considerable variation in approaches to drive impact. Computing exploits a broad range of knowledge transfer strategies and operates internationally to ensure it has maximum impact. We are concerned that the diversity of knowledge transfer strategies and techniques, and the international nature of computing research, are not sufficiently reflected in the current approaches to impact assessment.

A large proportion of computing impact is built upon an open approach to innovation, with systems made freely available in order to stimulate global innovation ecology. One high profile example of this is the open source software industry including products such as Linux or even

many parts of the world wide web. This approach to exploitation essentially relies on the efforts of others to drive exploitation. Para 68 says “ We do not envisage that a unit could claim credit for impact which was based on research undertaken in the unit but which was exploited or applied through the efforts of others, without a demonstrable contribution by the unit to that exploitation” This would essentially prohibit the reporting of this form of impact.

We agree that case studies are a good way of assessing impact and it will be necessary to rely on the assessors to judge the relevance and attribution of the case study and that indirect impact should also be allowed. We also feel that it is important that the criteria through which case studies are judged have some flexibility. The current approach appears to presume a particular small set of models of innovation and knowledge transfer. We would suggest that submitting units should be allowed to specify the particular strategy and approach to innovation and that sub-panels be allowed to judge these case studies against the discipline specific criteria best suited for that approach.

The lack of diversity of knowledge transfer activities is also reflected in terms of the descriptions of impact levels, which are currently described in terms of economic impact alone. We believe that broader societal and policy impacts should also be included.

Although outputs move with the people and count for the new institutions, we do not think that impact should move with people in general, but should be attributed to the institution where the work was originally undertaken. However there will be situations where impact is a result of collaborative research between multiple institutions and some situations where some of the work was done by a researcher at one institution but then additional related work continued at another institution after the researcher moved, so the impact will have to be attributed to more than one institution.

There is also impact that may result from multi-disciplinary work in a single institution which could legitimately be submitted to multiple units of assessment. With the increasing amount of collaborative work and interdependency of research ideas, the case studies will often end up relating to work done in many places. Given this, we believe it is essential that *any research* leading to impact is eligible.

We would like to highlight the significant overlap between the elements to be judged under impact and those constituting the engagement element of the environment in 79c. The key difference is time, with engagement being essential for fostering future paths to impact and ensuring the long-term sustainability of impact. We would argue that this suggests the need to ensure that environment is at least equally weighted to impact to avoid risking unhealthy behaviour, where effort focuses on short term impact at the cost of a long-term, sustainable contribution to the economic and social health of the nation.

We also feel that this overlap may lead to confusion. For example, we do not think that income from industry, government etc. should be included under impact. Having Research Council income under Environment and industrial income under Impact is confusing. EU funding for research in CS & I is a major, and highly competitive, source of research funding. Where does

EU income go? We are also concerned at the cost of changing accounting approaches in HEIs to allow this form of reporting. Income is a metric of *potential* to do research and is not a measure of research quality or impact so all income should be under environment.

Collaboration with users, in the form of industry, government and other disciplines, obviously impacts these 'users' of computing research but may not yet have a measurable economic or social impact. This type of impact could be assessed in terms of mini-case studies e.g. 100 word maximum descriptions of x case studies of collaborations where  $x = n/10$  for n FTE staff members.

In our experience, many of the esteem indicators returned in 2008 were really impact indicators e.g., membership of government, charity or industrial advisory boards; consultancies; leadership roles in industry, commerce, research councils, and learned societies.

The various types of indicators could have different weightings – we envisage higher weighting for the 500-word impact case studies than the 100-word collaboration case studies.

**Consultation question 4:** Do you have any comments on the proposed approach to assessing research environment?

There was no definition of the star levels for the environment sub-profile given in the consultation document, only for outputs and impact. We believe this is an oversight that needs to be rectified.

A detailed template of the indicators for environment is needed to aid assessment and give guidance to institutions preparing submissions. Although this may be subject specific, the following could be useful for many science and engineering subjects.

1. Infrastructure, facilities and administrative support for research
  - Existing infrastructure and facilities for research e.g. specific laboratories or programmer support for software tool development and evolution (200 words)
  - New institutional investment (200 words)
  - Administrative support in terms of generating and managing research proposals (100 words)
2. Translating Research into Impact
  - Institution/Departmental support to enable the translation e.g. innovation and consultancy support, training, exhibition or conference organisation support, time off for attending meetings. (300 words)
  - Specific training if appropriate. (100 words)
3. Arrangements for developing and supporting staff in their research

- Support provided for early career researchers with respect to teaching load, specific finance, training etc. (300 words)
  - Mentoring and management of research grant proposals e.g. internal reviewing, training (200 words)
  - Research team management/supervision training (100 words)
4. Research Student Training
- Research degrees (PhDs) awarded /FTE (Standard Analysis, RA3)
  - Current numbers of PhD Students
  - Research student training and support for conferences attendance etc. (300 words)
  - Arrangements for monitoring PhD students, encouraging completion (300 words)
5. Research income: funding strategy, amount received and sustainability
- Total income per FTE (Standard Analysis)
  - Composition of income
  - Trajectory of income
  - Explanation of income variations over assessment period (200 words)
  - Research Assistants per FTE (Standard Analysis)
  - Studentships per FTE (Standard Analysis)
  - Description of funding sources for studentships (200 words)
6. Credibility, vitality and sustainability of research organisation
- Free text: 100 words per FTE submitted
- Research Group structure and critical mass
  - Credibility, vitality and sustainability
  - Strategic vision
  - Leadership of discipline

**Consultation question 5:** Do you agree with our proposals for combining and weighting the output, impact and environment sub-profiles? If not please propose an alternative and explain why this is preferable.

As stated previously, we think the impact description is too narrow and the weighting too high. We need to recognise the whole pipeline of innovation from fundamental theoretical work to economic impact. In particular, it is very difficult for fundamental theoretical work to result in direct economic, social, public policy, cultural and quality of life impact. In many cases, the theoretical concepts influence the thinking and approach used by other academics who may

adapt the concepts into more practical and applied approaches. The current definition of impact explicitly rules out influence on scientific knowledge and academia, as this is meant to be covered by the outputs. For example, the Turing Award is computing's equivalent of the Nobel Prize. Hoare and Milner, at the time based in Oxford and Edinburgh, won their awards for work on foundational aspects of the discipline that influenced the work of other academics across the world on experimental programming languages, which in turn influenced the design of popular commercial programming languages like Java, and the technologies that underlie the internet. However this sort of impact may also happen a number of years after publication and in different institutions. Having 25% of the research weighting on a comparatively narrow scope of impact, which has never been previously assessed, could have a detrimental influence on fundamental high-quality research in the UK within CS& I as well as on subjects such as Mathematics. There is likely to be some management within institutions which will decide to cut back on any fundamental research as it will not rate highly on impact and focus mostly on applied research. Providing hard evidence of impact is likely to be very difficult in many cases and so many case studies are likely to contain much rhetoric. In fact, esteem is more tangible and easier to assess than impact. Too much emphasis on impact is likely to change the nature of the research being done in institutions and reduce quality.

Our view is that the weightings should be Outputs 60%, Environment 20% and Impact 20% particularly as Environment now includes esteem, so is a larger category than in RAE 2008.

**Consultation question 6:** What comments do you have on the panel configuration proposed at Annex E? Where suggesting alternative options for specific UOAs, please provide the reasons for this.

CS & I had 20 members for assessing 1846 FTEs and felt this was a burdensome workload for output assessment. Having an additional 12-15 people to help with specialised topics and impact assessment will help. However it is essential that anyone involved in output assessment takes part in the sub-panel initial calibration process otherwise their assessments will be useless.

This would also require classification of outputs into specific topics so that they can be passed to the relevant assessors at the start of the assessment. Sub-panels would have to define a set of topics for this classification.

Given the already very large size of the CS & I area, we would be opposed to any panel reorganisation that increased the size of the sub-panel covering CS & I. Moreover, CS & I research uses a wide diversity of methodologies and is often inherently interdisciplinary. Although CS & I researchers are used to assessing such research, researchers from other fields, within an enlarged REF panel, might have difficulty adjusting to such methodological diversity.

**Consultation question 7:** Do you agree with the proposed approach to ensuring consistency between panels?

Yes – we think the larger panels and uniform criteria will result in greater consistency.

**Consultation question 8:** Do you have any suggested additions or amendments to the list of nominating bodies? (If suggesting additional bodies, please provide their names and addresses and indicate how they are qualified to make nominations.)

**Consultation question 9:** Do you agree that our proposed approach will ensure that interdisciplinary research is assessed on an equal footing with other types of research? Are there further measures we should consider to ensure that this is the case and that our approach is well understood?

The CS & I sub-panel received considerable inter-disciplinary research in 2008 and this was assessed without problem so we do not envisage any problems for the REF.

Cross referrals have not worked in the past and should be discouraged due to the lack of consistency because of lack of calibration.

**Consultation question 10:** Do you agree that our proposals for encouraging and supporting researcher mobility will have a positive effect; and are there other measures that should be taken within the REF to this end?

Yes

**Consultation question 11:** Are there any further ways in which we could improve the measures to promote equalities and diversity?

We welcome the new proposals which provide an improvement in transparency and uniformity.

**Consultation question 12:** Do you have any comments about the proposed timetable?

We consider the timescales for the next REF to be unrealistic. There is not enough time for institutions to set up procedures for verifying citations, specifying impact and dealing with the changes in the information to be submitted. The assessment should be delayed by a year. This will give more time for evaluation of the proposed impact assessment pilot, give institutions more time to prepare changes in REF compared to RAE and give panels more time to define assessment criteria, which will have to substantially change from the 2008 RAE, as well as to define techniques for calibration of assessment, which will be essential for the larger sub-panels.

**Consultation question 13:** Are there any further areas in which we could reduce burden, without compromising the robustness of the process?

More effort is required in providing information to sub-panels in a timely fashion and in a format that will aid the assessment. In RAE 2008, CS & I subpanel members spent considerable time searching for missing or incorrect outputs, and processing data to get it into spreadsheets so that it could be easily assessed.

Having predefined detailed templates for Environment and Impact will help both institutions in preparing the information and panels in making assessment easier.

All outputs (journals and conference papers) should be provided, wherever possible, to HEFCE by HEIs in a machine-readable format, rather than HEFCE obtaining these centrally. This system caused considerable delays and frustration to the panel in RAE 2008. Often authors or publishers have pdf versions of books or reports and so many of these can also be provided by the institutions. The argument that if institutions provide the outputs, they are not verified is spurious. It is very unlikely that people will provide a 'better' output than one published and this is likely to be detected by the assessors. Random validation of outputs can be performed if this is deemed necessary.

The "sector-wide improvements in research information management" proposed in para 124 should supplement the huge institutional investment in such systems, rather than try to replace them, which would waste this investment. This can be done by defining a clear interface for institutional systems to interact with the REF central one.

**Consultation question 14:** Do you have any other comments on the proposals?

To summarise, the main points in our comments are:

- \* A broader definition of impact to include indirect impact, academic impact and impact on collaborators is essential
- \* Allow equal weighting of 20% for Impact and Environment as Impact is an untried form of research assessment and Environment now includes esteem. Too high a weighting for Impact is likely to have a detrimental affect on fundamental high-quality research.
- \* Delay the date of REF by at least one year, to allow evaluation of the trial Impact assessments, more time for sub-panels to prepare new evaluation criteria (particularly for Impact) and templates of required information, more time for institutions to prepare and validate required information, and sufficient time for assessors to take part in initial sub-panel calibration.
- \* CS&I sub-panel requires access to Google Scholar citations, in addition to any others that are provided, as they correlate more closely than WoS with the RAE 2008 assessment.
- \* Four outputs should be provided per person but sub-panels require sufficient assessors commensurate with the number of people being assessed and the number of assessors assigned to each output.
- \* More effort is required in providing information to sub-panels in a timely fashion and in a format that will aid the assessment.