

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. All responses should be e-mailed to 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.
 - 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.
 - c. Responses from institutions in Northern Ireland should be **copied to** the Department for Employment and Learning, e-mail research.branch@delni.gov.uk.
3. 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. Equivalent legislation exists in Scotland.

Respondent's details

Are you responding: (Delete one)	On behalf of an organisation
Name of responding organisation/individual	Society for the Study of Artificial Intelligence and Simulation of Behaviour (SSAISB or AISB)
Type of organisation (Delete those that are not applicable)	Academic association or learned society
Contact name	John Barnden (Prof.)
Position within organisation	Chair
Contact phone number	0121 414 3816
Contact e-mail address	j.a.barnden@cs.bham.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.

Impact: NO (ON MAJOR ASPECTS) – see Qn 3 answer. Main points: Some impact measurement may be reasonable, but it should not insist on departments etc. themselves converting research results into impact. **Two different things** should be distinguished and assessed: **having impactful research**; and **conversion** of results into impact. Our points are based not just on what appears to us logical, but also on (a) the pragmatics of assessment, and, very importantly, (b) properly serving the governmental policy aims that impact assessment is intended to serve in the first place.

Quantitative Indicators of Research such as citation information: We applaud the maintenance of expert review as the central tool. But we believe that the use of Qis needs to be explicitly constrained yet further. They should be **explicitly confined** to the assessment of “significance” of outputs, and indeed have only a limited role there. Further, “significance” outputs should be confined to **academic** significance. See answer to Qn 2.

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.

Citation info: We recognize that citation info may be useful, though only to aid and abet expert review. However, a more fundamental problem with it is that there are many reasons apart from research quality for a paper having many citations: e.g., it may survey others’ work in a convenient way; it may present a useful tool or methodology that does not itself rest on excellent or original research results developed by the authors; it may only describe application of already existing research; it may present an original but minor result that happens to be convenient in the development of more major results by other researchers. We suggest that what citation info could validly be held to contribute to measuring (though still approximately, unreliably, and partially) in the academic sphere is merely the **significance** of outputs. Thus, panels should be **explicitly restricted** to using quantitative indicators such as citations only in

their assessment of significance. However, this assessment should still be largely based on qualitative assessment, guided by the output commentaries provided by authors.

We are concerned that the rules as currently stated do not provide clear guidance to panels in how and to what extent they should use QIs, and do not ensure enough uniformity across panels. We would suggest that it is only exceptionally high citation counts that should influence a panel, this influence being to argue for an exceptionally high degree of significance. (What counts as exceptionally high will of course vary between panels.)

But panels should be required to assess *why* an output has high citation counts and not simply assume that a high count indicates real significance. It is often the case that a paper is widely cited because it provides a convenient summary of work in an area, rather than contributing significant ideas in its own right.

Moreover, “significance” as defined in para 39 needs to be *restricted to purely academic significance*, leaving economic/social/... significance to be covered, as it now should be, by the Impact component of the overall assessment. It is highly anomalous that significance as currently defined should include economic, social, etc. varieties. (Exceptions might be made in the case of some disciplines, e.g. practice-based social science, where academic significance cannot be untangled from non-academic impact.)

Citation counts might separately have a role to play, in some cases, in assessment of economic, social, etc. impact. This use of the counts should be carefully distinguished from their use in assessment of academic significance.

Under the above proposals, citation info is used not in a more defensible way but also in a much more manageable way (clearer and more definite rules can be drawn up, and uniformity across panels becomes easier to manage and more intrinsically justifiable).

The “star” categories (para 41): As in the case of the RAE, we feel that the introduction of the national/international distinction is confusing, unmanageable, multifarious and in any case unnecessary. A constant question in the RAEs has been what “international” amounts to. How can one ever know that a piece of work has only national recognition of rigour, etc. rather than international?? How international does something have to be to count as international? Are national standards of excellence different from international ones? We wonder what someone from another country must think on seeing the categories. Is the UK not confident that its “national” view of excellence is on a par with international standards?

The author of this response was intimately involved in the RAE 2008 preparation and submission process in his own department and across his university. He is not aware of any case where a distinction between 1* and 2* in the department and university’s own deliberations was actually drawn on the basis of a distinction between “national” and “international” standards.

We suggest that the definitions of the categories to be adjusted to something like the following, which excludes mention of [inter]nationality clearer and also clears up some other wording matters. (It can be separately explained, if it should be thought necessary at all, that the

standards to be applied should be ones that are internationally recognized across the discipline in question.) We have underlined the main loci of adjustment.

Four star: Exceptional: Research that meets the highest standards of excellence in terms of originality, academic significance and rigour. The significance needs to be at a world-leading level. [NB: it hardly makes sense to apply the epithet “world-leading” to originality and rigour.]

Three Star: Excellent: Research that meets high standards of excellence in terms of originality, academic significance and rigour, but that nonetheless does not meet the very highest standards of excellence.

Two Star: Good: Research that is good as regards originality, academic significance and rigour.

One Star: Acceptable: Research that merely meets basic standards of quality as regards originality, academic significance and rigour.

Unclassified: Research that falls below basic standards of quality, or work that does not meet the published definition of research for the purposes of the assessment.

3 versus 4 outputs: This seems like a good way of reducing work (for everyone), while not weakening standards *if* the REF were only four or so years away from now. But that is probably too soon anyway (see below).

Double-weighting of some outputs: This should be adopted, in a discipline-sensitive way. No rigid distinctions can be drawn, but it is mostly clear when, for example, a research monograph amounts to (at least) two normal research articles. Sub-panels should make the judgments on specific cases and as to the criteria that they will apply, and should make the criteria clear in their descriptions of working methods.

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.

In the following we put aside grave concerns we have about an excessive amount of weight being given to already-demonstrated [economic/social/...] impact. ***This could constrain research in such a way as to militate in the long term against the very sort of impact that the public would understandably like to see from academia.*** However, if impact is in some way to be included then:-

We applaud the careful thought that has gone into several aspects of the proposed scheme, and especially the use of case studies and avoiding expectations that every researcher should be engaged in impact generation. But ...

Paragraph 68 is exactly wrong in the view encapsulated in: "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 mixes up two things:*** having the research excellence to be able to generate ideas that turn out to have important and solid impact; and having the ability to convert research ideas (whether one's own or other people's) into impact. It seems to us entirely unjustifiable that the originators of an idea with major impact should fail to get credit just because they have not themselves done the exploitation. This is especially important point in view of contributions academics make to open-source software. More generally, the route to market often passes through a chain of academic and industrial researchers.

But conversely, we feel that conversion even of others' ideas into impact is worthwhile and should be credited.

Thus, the current scheme is doubly unfair and illogical! We strongly urge adoption of a scheme that separately measures both ***impactfulness*** – the extent to which a department or other unit's research has led to impact (however and wherever achieved) – and the extent to which the unit is engaged in ***conversion into impact*** (wherever the ideas came from).

This separation would have several pragmatic benefits for the assessment process, apart from being fairer and more logical. It would help to some extent with the ***time-lag*** problem of deciding how far back the exploited research ideas can go, as this issue would only be relevant to assessing impactfulness (it simply does not matter in conversion to impact how old the ideas are). It would help by encouraging separate sets of rules, criteria and indicators for impactfulness and conversion into impact, rather than mixing them up together.

Also, the current mixing-up of the two aspects is potentially dangerous to the very policy that impact assessment is meant to serve. Chasing conversion to impact could inhibit the production of the very sorts of ideas that could have long-lasting, sustainable, revolutionary, ... impact. Equally, requiring conversion to impact to be based on a unit's own ideas could inhibit the very sort of conversion activity the government would like to encourage.

We also wish to draw attention to two matters of general policy. First, whether an attempt at impact succeeds is often dependent on matters ***beyond the control of researchers*** or even

whole universities, making it invidious to compare different attempts without taking all relevant external factors into account, such as the politics of the day, the general economic climate, passing technological fashions, etc. Secondly, the impact assessment scheme in the consultation document falls into the trap that many have fallen into in the past of ***imagining that it is primarily the responsibility of universities to do conversion into impact***. Conversion is at most a responsibility shared with industry and government, and indeed is arguably more the responsibility of the latter two organs. Rather, ***it is impactfulness that universities might be expected to provide***.

A further glitch in the currently proposed scheme is that the “significance” component in research-output assessment allows for economic/social/... impact as well as academic impact/significance. This seems entirely illogical now that there is a separate impact arm of assessment. The impact part of a submission to REF could refer to relevant research outputs. Removing economic/social/... impact from research-output significance would also have the pragmatic benefit of easing the task of academics in assessing outputs and would reduce or eliminate the possible need for complex provisions for “users” to be involved in research-output assessment as opposed to impact assessment.

User involvement in impact assessment: We are worried by the statement in para 96(c) that it is mainly user members of panels who would assess impact submissions. Users should be strongly involved, but academics need also to be strongly involved, ***to ensure that the claimed impact is genuinely founded in research ideas*** rather than only having a superficial connection to a body of research. It is very easy for someone outside a discipline to have a wrong idea of the nature or significance of ideas within the discipline.

We are also concerned about how problems of ***users’ conflicts of interest*** will be adequately dealt with, given that users will be allied with particular sectors of the economy, etc. and with particular companies, etc. The REF scheme’s careful attention to such matters as the panel nomination process and to equality and diversity issues in returned-staff selection needs to be matched by proper safeguards concerning users.

Paragraph 72 [narrowness of impact]: We believe that this is wrongheaded in penalizing departments that exert impact only in a “[narrow] part of the territory that the unit might have been expected to cover”. The unit could well be able to be much more effective this way than attempting breadth. Surely what is important that university system as a whole cover the breadth of possible impact, rather than that each department do so. The provision in the paragraph could actually be ***detrimental to the success of government policy*** in deriving impact from the university system.

Paragraph 84 [uniformity of weighting]: We believe that this is wrongheaded in imposing (great) uniformity of weighting of outputs, impact and environment across different UoAs: The importance of impact, especially, could be *radically different* across different UoAs. Decisions on the issue should not be based on just simplicity of assessment. Also, we query why a university’s decision to place a unit in one UoA versus another on the basis of weighting differences should be pejoratively labelled as “tactical” – if the research in question can validly be assessed under two different UoAs, then it can validly be subject to different provisions for those UoAs.

Number of case studies: We have no problem with the suggestion of one case study per 5 to 10 submitted staff members, but we would urge that the range here should be used to cater for differences between different UoAs, rather than for there to be a single number used for all UoAs.

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

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.

60% for the actual research seems low. We would suggest 70% for research, 20% for impact (divided equally between impactfulness and conversion to impact) and 10% for research environment.

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.

[We are moved to comment on the following matter as Linguistics has important connections with Artificial Intelligence.]

It is illogical to put Linguistics under “European Languages and Studies”, as much of linguistics is importantly not about European languages. It would be even more illogical to include it under “English Language and Literature”. Also, the existing “European ...” category is already illogical and potentially controversial as it includes Latin American languages – bad enough in the case of Iberian-based languages, indefensible in the case of indigenous ones! Perhaps replace “European...” by “Linguistics and Foreign Languages and Studies”? The resulting scheme would still suffer from having Linguistics divorced from English L&L, but it would at least be better named.

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

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?

It cannot possibly be on an “equal” footing if the assessment is based on discipline-defined panels! The provisions in the proposed scheme are reasonable in themselves, but the broader panels will only help with certain limited types of interdisciplinarity. What about the sort of interdisciplinarity that we are concerned about in Artificial Intelligence, between Computer Science and areas such as Linguistics, Psychology and Philosophy? The general point here is not specific to AI: many other examples could be given of such “distant” links in other areas. Perhaps the intention is that specialist advisors and associate members still be used in this sort of case, but ***we would appreciate seeing an explicit statement in the final scheme about strong measures for handling “distant” types of interdisciplinarity.***

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?

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

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

2012 as a submission year seems pointlessly and unmanageably close. We endorse the comments on this question that are in the UKCRC response to the consultation.

Five years between REF exercises might be reasonable in steady state, but given the delay that the necessary consultations have introduced into devising the first one, 2012 is premature by at least two years.

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

Some of our suggestions in answer to Qns 2 and 3 would help reduce workload and complexity of work.

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