Stern Review Questions &
Joint Responses from BAFA/CPAF/CDAF

1. What changes to existing processes could more efficiently or more accurately assess the outputs, impacts and contexts of research in order to allocate QR? Should the definition of impact be broadened or refined? Is there scope for more or different use of metrics in any areas?

Currently the Higher Education Sector is in a period of transition with the proposals for a teaching excellence framework outlined in the Government Green Paper in November 2015 *Fulfilling Our Potential* with an increased emphasis on "A simpler system with students at the centre" (p. 57). The Green Paper is vague on the relationship between teaching and research in this new framework. We note that the Green Paper does register its commitment to the Haldane Principle and implied dual support of direct central government (QR) and research council funding (Para. 4, p. 69). We note that the Green Paper announced the intention of holding a new REF in 2021, and proposed reducing the cost of REF2021 through the use of metrics (p. 73), notwithstanding the general finding of the *Metric Tide* (2015) that peer review should remain a core component of future REFs.

BAFA/CDAF/CPAF have previously responded to the HEFCE Consultation on the use of metrics in June 2014, and expressed concerns that in the social sciences citation indices are neither fully comprehensive nor reliable, and argued that "on behalf of the accounting and finance community we strongly believe that peer review should remain the main method of assessing academic research outputs".

Some issues that arise from the current methods of measuring research quality are:

- With the setting of tuition fees in English universities in 2010 as an important component in the funding of higher education, reliance on QR funding has diminished in importance, and it is appropriate to question the need for a research assessment framework. (Although we note that regions within the UK have different funding structures). What is the purpose of the REF in a student-centred HE system?

- A major problem lies in the use of 'university quality' as a uni-dimensional concept measured by a position in a league table, that the REF and teaching metrics attempt to measure. Universities compete with respect to this single dimension by making decisions that will improve their position in the national league table. Such a system is full of perverse incentives (Bevan and Hood, 2006) and results in a system whereby what gets measured gets done. Is such a system appropriate for the development of the higher education sector? A wider, more multi-dimensional approach would be an improvement.

- Indicators of departmental research quality are biased due to the fact that only selected staff are entered. We believe that reliable measures should be developed that includes all academic staff (as in research assessments in other countries). At the very least submission rates for staff should be published along with REF results, and a requirement for the submission of all academic staff should be considered. Such a requirement would encourage institutions to provide a more enabling environment for research.

We fully support the objective of encouraging impact in research. However the current definitions for measuring impact raises issues, which we note below:

- We note that the current definitions of impact requires non-academics (users including policymakers, business people and others) who are willing to learn from academic research. Achieving impact requires users on the receiving end of research to be willing to be impacted, and in some areas of activity, users may not be receptive to research. For example in accounting, policy makers have not traditionally viewed research as particularly relevant to their work. As a consequence a researcher may do work that is very relevant to current practical issues in accounting, but it might have little impact. In finance, the Turner Report (2009) into the causes of the financial crisis blamed regulators for adhering to the principles of efficient markets, yet the main academic journals have argued consistently for the importance of behavioural finance and the existence of market inefficiencies for the previous 20 years.

- By rewarding 'impact' as currently interpreted, there is a danger of not rewarding researchers who are doing practically highly relevant work, but who have the misfortune to be working in areas where policy makers are not receptive to research. This seems perverse. Perhaps rewarding 'relevance' would be more to the point. Whether relevant research has impact as currently defined may be beyond the researcher's control;

- There is a concern that where impact is something that the researcher can influence, we may be assessing the researcher's impact rather than the impact of the research. Under the current REF rules, impact is recognised provided the cited research underlies the impact. However it is may be difficult to separate out the research's impact from the researcher's impact. And even when such separation is possible, we may wish to reward the researcher's engagement with practice. Under current definitions the process necessarily mixes up a quality of the research (its impact) with a quality of the researcher in engaging with practice. We should encourage both, but recognise that they are different;

- Impact is often interpreted as meaning "applied research." However, this may not be straightforward for certain subject areas, and a focus on applied research may neglect empirical and theoretical works that are important for generating a more general interpretation of impact in the longer-term. The current practice of 'Impact' cases focuses on the application of research outputs, but ignores the contribution of the output to the scholarly and theoretical development of both the subject and practice. We agree with the requirement that underlying research for the impact case needs to be of "international" quality, but would expect the system to recognise the many different forms of international research outputs;

- We note that much impact is generated through teaching in HE institutions, and that such impact should be reflected in either a teaching excellence framework or a research excellence framework.

2. If REF is mainly a tool to allocate QR at institutional level, what is the benefit of organising an exercise over as many Units of Assessment as in REF 2014, or in having returns linking outputs to particular investigators? Would there be advantages in reporting on some dimensions of the REF (e.g. impact and/or environment) at a more aggregate or institutional level? While the primary purpose of REF is QR resource allocation, data collected through the REF and results of REF assessments can also inform disciplinary, institutional and UK-wide decision making.

It should be recognised that journal articles in Accounting and Finance (the typical output type) tend to be substantive pieces of work of up to 50 pages, so that the number of outputs in consequence is small. This is a long-recognised feature of A&F research, and is similar in other countries. Zivney and Bertin (1992) examined the publishing record of 1,137 finance doctorates from 67 US universities awarded for the years 1963-1987. The subsequent publishing record of these graduates was tracked from 1969-1988 in 19 leading Finance Journals, 29 leading Accounting Journals and 80 additional top-ranked Business and Economics Journals. They conclude that "Our data indicate that publishing one article per year in any finance journal (or finance, accounting, economics or business journal) over any prolonged period of time is a truly remarkable feat." [p. 328]

Environment is an important factor for underlying research outputs and for generating research 'impact', and it is highly likely that research environment is positively correlated across units in the same institution. Therefore, measuring some components of research environment at the institutional level may be more appropriate.

We note that the REF2014 outputs on the HEFCE website are not attributable to the staff members at that university, so it would appear that HEFCE have chosen to not report returns linking outputs to particular investigators. The QR funding is allocated at an institutional level, and it is at the university's discretion whether to allocate funds to units in line with the REF scores. Our members report variable practices between universities on the links between resources and REF scores. If so, a more aggregate basis of assessment would be desirable. However there is a concern that smaller universities with only a limited number of centres of excellence, might suffer by aggregating at the institutional level.

We believe that the current level of aggregation with respect to subject areas within Business and Management is appropriate, although we note the large size of the B&M constituency (of which accounting and finance is a part) in REF2014 with 101 submissions and 3,320 FTE staff submitted. We note the use of sub-area reports in REF2014, and support their co0ntinued use.

3. What use is made of the information gathered through REF in decision making and strategic planning in your organisation? What information could be more useful? Does REF information duplicate or take priority over other management information?

REF information dictates resource allocation for research. In this way REF information informs management about priorities for resource allocation especially in the case of unfavourable REF outcomes. A concern is that it may substitute, possibly inappropriately for other information. That is, in the absence of REF information, universities would need to gather other and potentially more informative data as the basis for resource allocation.

4. What data should REF collect to be of greater support to Government and research funders in driving research excellence and productivity? The incentive effects of the REF shape academic behaviour, such as through the introduction of the impact criteria.

We have already expressed our support for the inclusion of impact as an indicator of research excellence, but have expressed concerns over the definitions of this indicator.

Our concerns are that as currently defined, the assessment of impact::

- Reduces cutting edge speculative research
- Focuses unduly on applied research that can easily be measured
- Does not allow for the varying propensity of users to engage with academic research
- Does not always distinguish between impact and public engagement
- Does not always distinguish between research impact and the impact of the individual researcher or research team (Esteem?)

Similar to the current format of 'impact' cases, a mechanism that can objectively assess the 'impact' of research output on academic development of the subject area would be desirable.

One measure of research output that is poorly measured currently is the quality of research output from PhD students. Currently, numbers of completing PhD students are a component of research environment. However this measure grossly understates the value of research supervision. A strong doctoral programme may generate research outputs for an early career researcher (ECR) that can be directly attributable to the research supervision, and yet under the current arrangements these research outputs would not be assigned to the research supervisor, but instead to the institution that employed the ECR. There is also a need to recognise that doctoral graduates may progress into senior non-academic jobs. A potential measure would be the "impact" of a doctoral research programme.

5. How might the REF be further refined or used by Government to incentivise constructive and creative behaviours such as promoting interdisciplinary research, collaboration between universities, and/or collaboration between universities and other public or private sector bodies? Previous studies have focused on the costs of REF with respect to the time and resources needed for the submission and assessment processes. The Review is also interested in views and any associated evidence that the REF influences, positively or negatively, the research and career choices of individuals, or the development of academic disciplines. It is also interested in views on how it might encourage institutions to `game-play' and thereby limit the aggregate value of the exercise.

The REF has markedly changed behaviour with regard to the type of research work being undertaken, particularly by junior academics.

- It has caused a concentration of effort on the production of journal articles in 'top' journals to the detriment of other forms of output and other types of research (e.g. more qualitative research and more applied work in the management area).
- It has homogenised research.
- It has supported the trend towards more formulaic and short-term research as against a broader development needed by the discipline.
- It has led to 'game-playing' around the selection of staff to be included and the inability of HEFCE to produce submission rate statistics at a unit level, and the use of part-time overseas-based faculty hired on the basis of REF returnable outputs
- The REF needs to be a far lighter touch and more restricted in the purposes it seeks to achieve.
- While the REF has provided incentives for academics to co-operate with research teams at other universities, but it has created disincentives to work with colleagues at the same institution. Although REF2014 allowed up to two submission of the same output, this required a statement justifying the submission, and risk-averse institutions still appear to be nervous about submitting such work. Might it not be possible to remove this disincentive? Perhaps allocated percentages based on the number of authors?
- Because of the transferability of outputs across institutions, it has distorted the labour market for academics. Such distortions could be reduced by requiring outputs to be shared across institutions in the case of academics who move between universities.

With respect to interdisciplinary research, potentially such research can help generate research output that has better impact on businesses and policies. This could be encouraged at institutional level by offering further incentives for collaborative research between academic units and the users of research eg: professionals, businesses, government units, policy makers etc.

6. In your view how does the REF process influence, positively or negatively, the choices of individual researchers and / or higher education institutions? What are the reasons for this and what are the effects? How do such effects of the REF compare with effects of other drivers in the system (e.g. success for individuals in international career markets, or for universities in global rankings)? What suggestions would you have to restrict gaming the system?

There is evidence of a dramatic improvement in the quality of research undertaken at UK universities between the two periods 2001-2007 and 2008-2013. BiS (2015) noted that across all disciplines "The results of REF2014 showed significant improvements in the quality of research produced by UK universities since 2008: with 72% of UK universities' research outputs now assessed at the highest international levels compared with 51% in 2008" (p. 70) . More specifically within business and management, Pidd and Broadbent (2015) report the substantial increase in 4-star outputs in B&M relative to the two previous sub-panels in RAE2008 in B&M and A&F. In Table 2 of their paper, they report that the percentage of outputs graded 4-star increased to 20.5% (from 13.8% and 5.9% respectively) and suggest a number of factors explaining this improvement including "better and more ambitious research, an increased emphasis on internationalisation, greater receptiveness of international journals to European research and increased selectivity in submissions when compared to RAE 2008”. It should also be noted that there were rule changes from RAE2008. In REF2014 institutions were allowed to submit up to four papers from individuals who were on part-time contracts, and many institutions submitted such individuals. Another rule change was that the same institution was allowed to submit the same output from up to two individuals, whereas previously a high quality output could only have been submitted on behalf of one individual.

We recognise that research outputs are not only used for assessing universities' research but also used for individuals' career development, with most universities relying on references from senior academics on the quality of research outputs. Beattie and Goodacre (2012) investigate publication profiles of 137 accounting and finance faculty promoted to professor at UK universities during 1992–2007. On average, nine papers in established academic journals, with 5 at the highest 3*/4* quality levels in a portfolio of 20 outputs are required for promotion. Multivariate models of key performance benchmarks (quality and quantity measures) are constructed and have good explanatory power. They observe that publication requirements seem to have increased over time, mainly attributable to government-initiated Research Assessment Exercises. For internal promotions, there is some evidence of higher hurdles but no evidence that quality requirements differ based on gender; sub-discipline; research intensity of institution peer group; or government-initiated research ranking of unit.

Current practices drive a range of undesirable behaviours. These have helped drive the current trend to the variety of league tables that actually have little value in rational decision-making for most people:

- The REF and NSS (National Student Survey) scoring means that HEIs now strive towards a single objective of improving their league table positions and generally with strategies that are virtually indistinguishable from each other. A more vibrant university system would have institutions specialising in their niche areas where they are truly valuable and different rather than the unified model of a common approach with universities being distinguished only by their success in achieving these narrow aims.

- The conformity of research and lack of intellectual curiosity reduces the attractiveness of an academic career to the brightest and best students. Such students may move outside the HE environment;

- There is a concern that the requirement on young ECRs to produce a high quality output per year places undue stress on new academics, which may act as a deterrent to a career in academia;

- The current system appears to drive mechanistic styles of management at all levels in universities (departmental, faculty and central) which have been shown to be counter-productive in many respects (Franco-Santos et al, 2014). A more multi-dimensional system less subject to mechanistic modelling by universities would be more desirable.

7. In your view how does the REF process influence the development of academic disciplines or impact upon other areas of scholarly activity relative to other factors? What changes would create or sustain positive influences in the future? Much of REF focuses on the retrospective analysis of success achieved by institutions either through output or impact. Yet the resources provided anticipate continued success based on that track record. Are there means of better addressing forward-looking institutional plans and priorities, and how these might feed in to national policy?

The forward looking aspect is dealt with by the research councils. But universities still need some of discretionary funding they can use to support distinctive individual strategies. A process whereby a sum of discretionary funding is given to a university, perhaps subject to it submitting its plans on how it will be used would be a better mechanism for promoting innovation and diversity in research.

Some research areas take time to develop, and having a REF cycle within a specific research window can affect the commitment to developing new and risky research areas. Such a feature of research time-lines may also have implications for the outcomes based on impact.

8. How can the REF better address the future plans of institutions and how they will utilise QR funding obtained through the exercise? The Review is keen to hear of creative ideas and insights and to be open in its approach.
A 5 yearly report of the funds given to the institution, and the ways they have been used, with an indication of the outcomes obtained would surely be adequate. In the disciplines that require little equipment, the main investment would be in providing staff time to engage in research. The issues involved in funding the sciences with heavy equipment expenses and required capital investments are very different and probably require a different funding mechanism.

9. Are there additional issues you would like to bring to the attention of the Review?
Business schools have their own distinct research missions and environments in which these goals are pursued. The current system homogenises research strategies and serves to deflect them from their core mission and needs to be radically revised if they are to sustain diversity and maintain their value to society and enhance sustainable economic growth. Note.

While supporting the principles of encouraging impactful research, the current rules for attributing impact require three conditions: a) underlying pieces of academic research; b) individuals who conducted them, and c) institutions at which they were employed at specific points in time. These three conditions are unhelpful in encouraging all academics and universities to bring the results of all research (wherever conducted) to bear in a variety of ways that have benefits to the UK economy and social services.

John Cullen (BAFA)
Gillian Holmes (CDAF)
Ian Tonks (CPAF
23rd March 2016

References

Beattie, V. and A. Goodacre (2012). "Publication records of accounting and finance faculty promoted to professor: evidence from the UK." *Accounting and Business Research*, 42(2): 197-231.

Bevan, Gwyn, and Christopher Hood (2006) "What's measured is what matters: targets and gaming in the English public health care system." *Public administration* 84.3: 517-538.

Franco-Santos, M., Rivera, P., and Bourne, M. (2014) "*Performance Management in UK Higher Education Institutions: The Need for a Hybrid Approach"*, Leadership Foundation for Higher Education, London.

Business Innovation and Skills, Department of, (2015) "*Fulfilling our potential: Teaching Excellence, Social Mobility and Student Choice*", November, Cm 9141 (BiS, London).

Pidd, Michael, and Jane Broadbent (2015) "Business and Management Studies in the 2014 Research Excellence Framework." *British Journal of Management* 26.4: 569-581.

Turner, A. (2009), *A regulatory response to the global banking crisis*, March (Financial Services Authority).

Wilsdon, James (2015) *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*, July 2015, (HEFCE, London).

Zivney, T. L. and W. J. Bertin (1992). "Publish or perish: What the competition is really doing." *Journal of Finance*, 47(1): 295-329.